(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/44994 A2

(51) International Patent Classification[7]: G06F 19/00

(21) International Application Number: PCT/US01/45705

(22) International Filing Date: 30 November 2001 (30.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/250,112	30 November 2000 (30.11.2000)	US
60/250,449	30 November 2000 (30.11.2000)	US
09/771,332	26 January 2001 (26.01.2001)	US
09/782,702	13 February 2001 (13.02.2001)	US
60/285,895	23 April 2001 (23.04.2001)	US
60/288,229	2 May 2001 (02.05.2001)	US
60/289,764	9 May 2001 (09.05.2001)	US
60/304,521	11 July 2001 (11.07.2001)	US
60/307,660	25 July 2001 (25.07.2001)	US
09/915,063	25 July 2001 (25.07.2001)	US
60/308,878	31 July 2001 (31.07.2001)	US
60/311,582	10 August 2001 (10.08.2001)	US
09/929,135	14 August 2001 (14.08.2001)	US
09/930,535	15 August 2001 (15.08.2001)	US
09/930,688	15 August 2001 (15.08.2001)	US
09/930,646	15 August 2001 (15.08.2001)	US
09/930,543	15 August 2001 (15.08.2001)	US
60/326,549	2 October 2001 (02.10.2001)	US
60/238,312	10 October 2001 (10.10.2001)	US
60/329,113	12 October 2001 (12.10.2001)	US
60/328,861	12 October 2001 (12.10.2001)	US
Not furnished	19 October 2001 (19.10.2001)	US
Not furnished	26 October 2001 (26.10.2001)	US
Not furnished	13 November 2001 (13.11.2001)	US
Not furnished	30 November 2001 (30.11.2001)	US

(71) Applicant *(for all designated States except US)*: **THIRD WAVE TECHNOLOGIES, INC.** [US/US]; 502 South Rosa Road, Madison, WI 53719 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BROWER, Amy** [US/US]; 8519 Elderberry Road, Madison, WI 53717 (US). **BROW, Mary, Ann** [US/US]; 910 Pebble Beach Drive, Madison, WI 53717 (US). **CRACAUER, Raymond, F.** [US/US]; 3108 Creekview Drive #6, Madison, WI 53562 (US). **FORS, Lance** [US/US]; Hidden Hollow Trails #12, Madison, WI 53719 (US). **GRANSKE, Rocky** [US/US]; 2933 Manchester Road, Madison, WI 53719 (US). **DE ARRUDA INDIG, Monika** [US/US]; 6618 Montclair Lane, Madison, WI 53711 (US). **KURENSKY, David** [US/US]; 3279 N. Cramer, Milwaukee, WI 53211 (US). **LUEDTKE, Craig** [US/US]; 305 Edgemere Court, Waunakee, WI 53211 (US). **LUKOWIAK, Andrew, A.** [US/US]; 2254 High Ridge Trail, Madison, WI 53713 (US). **LYAMICHEV, Victor** [RU/US]; 2523 Carriedale Court, Madison, WI 53711 (US). **NERI, Bruce, P.** [US/US]; 5714 Kilkenny Place, Madison, WI 53711 (US). **REIMER, Ned, D.** [US/US]; 7125 Gladstone Drive, Madison, WI 53219 (US). **ROEVEN, Robert, T., P.** [US/US]; 324 Lowell Street, Stoughton, WI 53589 (US). **SKRZYPCZYNSKI, Zbiginiev** [PL/US]; 881 Orchid Street, Verona, WI 53711 (US). **ZIARNO, Witold, A.** [US/US]; 6301 Offshore Drive #319, Madison, WI 53705 (US). **COMERFORD, John** [US/US]; 7755 Summerfield Drive, Vernona, WI 53593 (US). **STUMP, Steven** [US/US]; 575 Harvest Lane, Verona, WI 53593 (US). **VIEGUT, Daniel, D.** [US/US]; 17 Jacobs Court, Madison, WI 53711 (US).

(74) Agents: **CASIMIR, David, A.** et al.; Medlen & Carroll, LLP, Suite 350, 101 Howard Street, San Francisco, CA 94105 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) **Title:** SYSTEMS AND METHODS FOR DETECTION ASSAY ORDERING, DESIGN, PRODUCTION, INVENTORY, SALES AND ANALYSIS FOR USE WITH OR IN A PRODUCTION FACILITY

(57) **Abstract:** The present invention relates to detection assay development, production and optimization. In particular, the present invention provides systems and methods for acquiring and analyzing biological information. The present invention also provides detection assay production with improved oligonucleotide synthesis and processing systems. The present invention further provides systems that integrate biological information collection with detection assay production that allow for rapid development of commercial products, such as analyte specific reageants (ASRs) and in vitro diagnostics (IVDs).

WO 02/44994 A2

Published:

— *without international search report and to be republished upon receipt of that report*

SYSTEMS AND METHODS FOR DETECTION ASSAY ORDERING, DESIGN, PRODUCTION, INVENTORY, SALES AND ANALYSIS FOR USE WITH OR IN A PRODUCTION FACILITY

The present Application claims priority to U.S. Application Serial No. not yet assigned, filed on November 30, 2001, with express mail number EL837033437US, herein incorporated by reference.

The present Application claims priority to U.S. Provisional Application 60/328,312 filed October 10, 2001, hereby incorporated by reference in its entirety by reference.

The present Application claims priority to U.S. Provisional Application 60/288,229 filed May 2, 2001, hereby incorporated by reference in its entirety.

The present Application claims priority to U.S. Provisional Application 60/329,113 filed October 12, 2001, and U.S. Provisional Application Serial No. not yet assigned filed October 19, 2001 with express mail number EL837033128US, both of which are hereby incorporated by reference in their entirety.

The present Application claims priority to U.S. Provisional Application 60/250,449 filed November 30, 2000; U.S. Provisional Application 60/250,112 filed November 30, 2000; and U.S. Provisional Application 60/285,895 filed April 23, 2001, all of which are hereby incorporated by reference in their entirety.

The present Application is a continuation-in-part of U.S. Application Serial No. not yet assigned, filed on November 13, 2001 with express mail number EL790816213US, which is a continuation-in-part of U.S. Application Serial No. not yet assigned, filed on October 26, 2001 with express mail number EL790816244US, which is a continuation-in-part of U.S. Application 09/782,702 filed February 13, 2001, which is a continuation-in-part of U.S. Application 09/771,332 filed January 26, 2001, all of which are hereby incorporated by reference in their entirety.

The present Application is a continuation-in-part of U.S. Applications 09/930,543; 09/930,646; 09/930,688; and 09/930,535 all filed on August 15, 2001, and all of which are hereby incorporated by reference in their entirety.

The present Application claims priority to U.S. Provisional Application Serial Number 60/289,764 filed May 9, 2001, and U.S. Provisional Application 60/326,548, filed October 2, 2001, both of which are hereby incorporated by reference in their entirety.

The present Application claims priority to U.S. Provisional Application 60/311,582 filed August 10, 2001; U.S. Provisional Application 60/308,878 filed July 31, 2001; and U.S. Provisional Application 60/307,660 filed July 25, 2001, all of which are hereby incorporated by reference in their entirety.

The present Application is a continuation-in-part of U.S. Application 09/929,135 filed August 14, 2001, and is hereby incorporated by reference in its entirety.

The present Application is a continuation-in-part of U.S. Application 09/915,063 filed July 25, 2001, which claims priority to U.S. Provisional Application 60/304,521 filed July 11, 2001, both of which are hereby incorporated by reference in their entirety.

The present Application claims priority to U.S. Provisional Application 60/328,861 filed October 12, 2001, and is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to detection assay ordering, development, production, sales and optimization methods and systems for the commercialization of products, such as research use only products (RUOs), analyte specific reagents (ASRs) and in vitro diagnostics (IVDs).

BACKGROUND

As the Human Genome Project nears completion and the volume of genetic sequence information available increases, genomics research and subsequent drug design efforts increase as well. There exists a need for systems and methods that allow for the efficient ordering, development, production and sales of detection assays that can be used in genomics research, drug design, and personalized medicine. A number of institutions are actively mining the available genetic sequence information to identify correlations between genes, gene expression and phenotypes (e.g., disease states, metabolic responses, and the like). These analyses include an attempt to characterize the effect of gene mutations and genetic and gene expression heterogeneity in individuals and populations. However, despite the wealth of sequence

information available, information on the frequency and clinical relevance of many polymorphisms and other variations has yet to be obtained and validated. For example, the human reference sequences used in current genome sequencing efforts do not represent an exact match for any one person's genome. In the Human Genome Project (HGP), researchers collected blood (female) or sperm (male) samples from a large number of donors. However, only a few samples were processed as DNA resources, and the source names are protected so neither donors nor scientists know whose DNA is being sequenced. The human genome sequence generated by the private genomics company Celera was based on DNA samples collected from five donors who identified themselves as Hispanic, Asian, Caucasian, or African-American. The small number of human samples used to generate the reference sequences does not reflect the genetic diversity among population groups and individuals. Attempts to analyze individuals based on the genome sequence information will often fail. For example, many genetic detection assays are based on the hybridization of probe oligonucleotides to a target region on genomic DNA or mRNA. Probes generated based on the reference sequences will often fail (e.g., fail to hybridize properly, fail to properly characterize the sequence at specific position of the target) because the target sequence for many individuals differs from the reference sequence. Differences may be on an individual-by-individual basis, but many follow regional population patterns (e.g., many correlate highly to race, ethnicity, geographic local, age, environmental exposure, etc.). With the limited utility of information currently available, the art is in need of systems and methods that can optionally be used in one or more production facilities for acquiring, analyzing, storing, and applying large volumes of genetic information with the goal of providing an array of one or more types of detection assay technologies for research and clinical analysis of biological samples. It is an object of the invention to fill these various needs.

SUMMARY OF THE INVENTION

The invention provides systems and methods for ordering, manufacturing and selling detection assays, and instrumentation related thereto. The system includes one or more components, such as a computer-based customer order component for ordering at least one of a plurality of oligonucleotide detection assays, and/or related instrumentation; a detection assay production component for creating the oligonucleotide detection assays; a shipping component

for shipping said oligonucleotide detection assays and/or related instrumentation; and a billing component for billing a customer for the oligonucleotide detection assays and/or related instrumentation. Optionally, the billing component comprises a payment receipt component for receiving payment for the oligonucleotide detection assays.

The present invention further provides systems, methods, and compositions that provide comprehensive solutions for the manufacturing, use, analysis, and sales of detection assays (e.g., oligonucleotide detection assays). For example, the present invention provides systems and methods for the ordering of detection assay, including electronic ordering (e.g., over public or private electronic communication networks) by general customers, as well as, distributors, collaborators, health care professionals, individuals, and established long-term customers. The present invention also provides systems and methods for detection assay design, including electronic quality assessment methods of detection assay components and design of primers (e.g., amplification primers) and probes. Assay design is made possible for large numbers of diverse assays (of a single type or of multiple types) and for large-scale production thereof, including the design of panels, research products, and clinical products (e.g., in vitro diagnostic products). The present invention also provides systems and methods for detection assay production, including coordinated synthesis, preparation, and quality control of detection assay components, and also detection assay assembly on a variety of presentation platforms, including 96, 384, 1536 well plates, and combinations thereof, slides, and other presentation platforms. Inventory control systems and methods, and design and production management systems and methods, are also provided for complete detection assays, for detection assay components, reagents for the creation of detection assays, and instrumentation used to manufacture detection assays. The present invention also provides systems and methods for selling detection assays, and systems and methods for assisting detection assay users in the collection and analysis of data produced by the use of the detection assays (of a single variety or of multiple varieties). The present invention also provides systems and methods for collecting, analyzing, and storing data, including detection assay design data and data generated by the use of the detection assays. Each of the components of the systems and methods of the present invention may be integrated to provide comprehensive systems and methods for the manufacture and use of detection assays, with exchange of data between various components of the system to optimize utilization of the

data generated by the detection assay or detection assay usage. Integration provides, by way of further example, methods to coordinate the movement of genetic information from research applications to in vitro diagnostic applications. Each of the components of the present invention are described in detail below.

In some embodiments, the computer based customer order entry component further comprises a consumer direct web order entry component. Consumers, include by way of example, the purchasing public. The computer based customer order entry component further includes home or work computers, workstations, PDAs or web appliances of members of the public. In other embodiments, the computer-based customer order entry component provides a unidirectional, bi-directional or omni-directional data feed into the detection assay production component, other components of the system and/or portions thereof. In certain embodiments, the data feed affects production cycles of the oligonucleotide detection assays. In particular embodiments, the data feed comprises statistical information associated with or related to one or more oligonucleotide detection assays of a single variety or one or more oligonucleotide detection assays of one or more varieties. In other embodiments, the statistical information is selected from the group consisting of total oligonucleotide detection assays ordered or oligonucleotide detection assay orders received; a histogram; an oligonucleotide detection assay average per consumer; an arithmetic mean; quantity of oligonucleotide detection assays, size of order of oligonucleotide detection assays; format of panel information; a mode; a median; a weighted mean; a harmonic mean; a geometric mean; a logarithmic mean; a root mean square; a root sum square, and combination thereof; a normal distribution curve, the normal distribution curve includes, but is not limited to, a normal distribution curve of number of consumers, number of detection assays, quantity of oligonucleotide detection assays, quantity of oligonucleotide detection assays or a certain type; a spread; a variance; a standard deviation; a skewed distribution; a sampling; a confidence level; and, a regression analysis.

In some embodiments, the present invention provides a system and method for manufacturing and selling detection assays, comprising one or more of the following components: a computer-based customer order component for ordering at least one of a plurality of oligonucleotide detection assays; a detection assay production component for creating the oligonucleotide detection assays of one or more varieties; a shipping component for shipping the

oligonucleotide detection assays; and a billing component for billing a customer for the oligonucleotide detection assays. In some embodiments, the billing component comprises a payment receipt component for receiving payment for the oligonucleotide detection assays.

In some embodiments, the computer-based customer order component comprises a client-based computer network, a physician's computer network, and insurance company computer network, a health maintenance organizations computer network, a hospital computer network, a distributor-based computer network, and/or a combination thereof. In some preferred embodiments, the computer-based customer order component comprises a web-based user interface for ordering the oligonucleotide detection assay via single or multiple linked screens or web pages. In some preferred embodiments, the web-based user interface provides a detection assay locator component. For example, in some embodiments, the detection assay locator component comprises a library of detection assay data from which an oligonucleotide detection assay can be selected from a single type of detection assays or from a catalogue of different types of detection assays. In some preferred embodiments, the library of detection assay data comprises single nucleotide polymorphism ("SNP") data or other data related to the SNP data.

In some embodiments, the detection assay production component comprises a shop floor control system (e.g. comprising an oligonucleotide control system for synthesizing oligonucleotides, and a centralized control network for processing oligonucleotides). In some embodiments, the shop floor control system is configured to direct oligonucleotide detection assay production using a make-to-order routine, a make-to-stock routine, and/or a fulfill-from-stock routine, or other software package. In some embodiments, the shop floor control system comprises a library of detection assay data from which the plurality of detection assays of a single variety or detection assays of more than one variety can be created. It is appreciated that this library of data, the accuracy of which has been checked against a single or plurality of databases of this type of data reduces the error rates associated with detection assay production.

In certain embodiments, the order entry component or the billing component comprises a differential pricing component. The differential pricing component is a set of routines that run on one or more processors of the system described herein. In other embodiments, the differential pricing component is capable of selectably pricing a detection assay or a single variety or a plurality of detection assays of more than one variety based upon a predetermined category of

product. In some embodiments, the predetermined category of product is selected from the group consisting of an RUO product, an ASR product, and an IVD product. These routines analyze the product category selection of a consumer or other purchaser to correlate the correct pricing for a detection assay with the category selected by the consumer or the end user. In additional embodiments, the differential pricing component comprises a routine that associates a predetermined price of a detection assay based upon a presentation platform selection. For example, if a consumer selects a 96 well plate as the detection assay presentation platform one price data set is correlated with the transaction. If the consumer selects a combination of different presentation platforms, e.g. 1536 well format, and glass slide format the routines correlate and tabulate the correct price data for the transaction.

In some embodiments, the detection assay production component comprises a synthesis component, a cleave/deprotect component, a purification component, a dilute and fill component, and/or a quality control component. In some embodiments, the synthesis component comprises a plurality of oligonucleotide synthesizers or a single synthesizer capable of a multiplicity of syntheses. The present invention is not limited by the nature of the synthesizers. Synthesizers include, but are not limited to, alone or in combination, MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), OligoPilot (Amersham Pharmacia,), the 3900 and 3948 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA), POLYPLEX (Genemachines), 8909 EXPEDITE, Blue Hedgehog (Metabio), MerMade (BioAutomation, Plano, Texas), Polygen (Distribio, France), and PrimerStation 960 (Intelligent Bio-Instruments, Cambridge, MA). Other synthesizers used herein are those that are capable of simultaneously creating 384 wells and 1536 wells of oligonucleotides. In some embodiments, the detection assay production component comprises an inventory control component. The inventory control component comprises hardware, software, an optional freezer or cooler (walk in style cooler in one variant) with selectable temperature control, and robotics to place and select items of inventory in predetermined locations within the freezer, cooler or cold room.

The present invention is not limited by the nature of the detection assay. In some embodiments, the detection assay comprises an invasive cleavage assay, a TAQMAN assay, a sequencing assay, a polymerase chain reaction assay, a hybridization assay, a hybridization assay employing a probe complementary to a mutation, a microarray assay (e.g. on a solid support), a

bead array assay, a primer extension assay, an enzyme mismatch cleavage assay, a branched hybridization assay, a rolling circle replication assay, a NASBA assay, a molecular beacon assay, a cycling probe assay, a ligase chain reaction assay, and a sandwich hybridization assay. In some embodiments, the detection assay is configured to detect a sequence selected comprising a polymorphism, a transgene, a splice junction, a mammalian sequence, a prokaryotic sequence, and a plant sequence. It is appreciated that one or more of these detection assays can be produced in one or more production facilities using the systems and methods of the present invention. Moreover, one ore more of these detection assays have data associated or related to each respective detection assay presented via the detection assay locator. By way of further example a particular location on the detection assay locator web page or screen can have listings for several types of detection assay for a single nucleotide polymorphism including pricing information for each respective detection assay. Moreover, it is appreciated that the pricing data located thereon can be variable. For example, where there are three types of detection assay on a page, a routine automatically makes pricing for a favored or predetermined detection assay lower or competitive with one or more other types of detection assays.

In some embodiments, the detection assay production component comprises an oligonucleotide detection assay design component. In some preferred embodiments, the detection assay design component comprises a PCR primer creation component that can optionally be used alone or in combination with the detection assay design component. In some embodiments, the PCR primer creation component is configured to optimize PCR primer concentrations. In some embodiments, the detection assay design component is configured to design a single type of detection assay, a plurality of detections assays of a single variety, or a plurality of detection assays or multiple varieties for detecting the presence of one or more polymorphisms (e.g., single nucleotide polymorphisms), RNA, other sequences and/or combinations thereof. In some embodiments, the detection assay design component is configured to design a panel or array comprising a plurality of oligonucleotide detection assays of a single variety, of multiple varieties, for a single SNP, for multiple SNPS, for a single SNP detected by multiple varieties of detection assays, and for multiple SNPs detected by multiple varieties of detection assays. In some preferred embodiments, the detection assay production component comprises a genotyping component. In some embodiments, the genotyping

component is configured to test an oligonucleotide detection assay (of a single type or multiple types) against a plurality of target sequences from different sources.

In some embodiments, the present invention provides detection assay ordering systems, comprising a first processor (including one or more microprocessors) in electronic communication with: a) a computer system or single computer of a customer; b) an electronic detection assay identification catalogue going across one or more genomic landscapes; c) a second processor (including one or more microprocessors) configured to carry out detection assay design; and d) a third processor (including one or more microprocessors) configured to carry out detection assay production. It is appreciated that processors one through three can be a single processor or multiple processors located in one or more locations. Moreover, it is appreciated that archival backup routines and devices provide back up for the data and routines used on one or more devices and components described herein. In some embodiments, the detection assay comprises an invasive cleavage assay or other assay described herein. In other embodiments, the first processor provides a user interface to the computer system of the customer. In particular embodiments, the user interface comprises stacked databases, or linked web pages. In further embodiments, the stacked databases, screens or web pages comprise SNP data or sequence data that includes a SNP. In certain embodiments, the stacked databases or web pages comprise pre-existing detection assay data. In some embodiments, the pre-existing detection assay comprises data of a detection assay that has passed through an in silico process. In particular embodiments, the pre-existing detection assay data comprises data of a detection assay that has passed through a genotyping process.

The present invention provides systems and methods for acquiring and analyzing biological information obtained from the use of one or more types or varieties of detection assays ordered or produced using the systems and methods described herein. For example, the present invention provides systems and methods for the use of genetic information in the generation of assays for detecting the genetic identity of samples, the production of assays, the use of assays for gathering genetic information of individuals and populations, and the storage, analysis, and use of the obtained information.

For example, the present invention provides a method for screening candidate oligonucleotides for use in a detection assay, comprising, providing 1) a candidate

oligonucleotide, 2) five or more target nucleic acids (e.g., 6, 7, 8, . . ., 100, . . .), wherein each of the five or more target nucleic acids is derived from a different subject; and detection assay components that permit detection of the target nucleic acids in the presence of a functional detection oligonucleotide; treating together the five or more target nucleic acids with the candidate oligonucleotide in the presence of the detection assay components; and determining if the candidate oligonucleotide is a functional detection oligonucleotide for use with each of the five or more target nucleic acids. In some embodiments, the target nucleic acids comprise a single nucleotide polymorphism. In some embodiments, the candidate oligonucleotide comprises a hybridization probe. In some preferred embodiments, the candidate oligonucleotide is designed to hybridize to a target sequence of at least one of the target nucleic acids. In some embodiments, the target sequence is identified by or selected by in silico analysis. In certain particular embodiments, the detection assay components comprise detections assay components for performing an INVADER assay. In some embodiments, the method further comprises the step of preparing a kit containing the candidate oligonucleotide if the candidate oligonucleotide is determined to be a functional detection oligonucleotide. In some embodiments, the kit comprises instructions, directing a user of the kit to use the kit with samples from subjects suspected of possessing any of the target nucleic acids from which the candidate oligonucleotide was determined to be a functional detection oligonucleotide.

The present invention also provides a method of gathering and storing genomic data derived from a detection assay, comprising providing a detection assay configured to detect the presence or absence of a nucleic acid sequence in a sample; a first computer system comprising one or more computer processors and a computer memory; a second computer system comprising one or more computer processors and computer memory, wherein the computer memory comprises a genomic information database; and a test sample; treating the test sample with the detection assay to generate test result data; collecting the test result data with the first computer system; and transmitting the test result data from the first computer system to the second computer system under conditions such that the test result data is added to the genomic information database of the second computer system. In some embodiments, the detection assay comprises assays including, but not limited to, hybridization assays, cleavage assays, amplification assays, sequencing assays, and ligation assays. In some preferred embodiments,

the detection comprises an INVADER assay, a TAQMAN assay, any other type of assay described herein, and/or combinations thereof. In some embodiments, the nucleic acid sequence comprises a single nucleotide polymorphism or RNA. In some preferred embodiments, the first computer system or computer including a microprocessor comprises one more detectors (e.g., fluorescent detectors, luminescent detectors, optical detectors, and radioactivity detectors). It is appreciated that the instrumentation described herein can also be sold as kit which would include the instrumentation described herein as well as a plurality of pre-ordered or ordered detection assays. In some embodiments, the test sample comprises a genomic DNA or RNA sample or a synthetic DNA or RNA sample. In other embodiments, the test sample comprises an RNA sample, and/or a PCR target/sample. In some embodiments, the test result data comprise information related to a subject from which the test sample was derived. Test result data can be presented to a user via a computer or workstation communicatively linked to any computer or display linked to any of the components described herein. In some embodiments, the first computer system (which is optionally networked) or computer is located in a different geographic location from the second computer system (which is optionally networked in a LAN, MAN, WAN, or combination thereof) or computer. In some embodiments, the transmitting comprises sending the test result data over a communication network on which the various computers are communicatively linked. In some preferred embodiments, the test result data comprises allele frequency information. In other preferred embodiments, the genomic information database comprises database data comprising allele frequency information, genetic location pathway data, metabolic pathway data, and/or combinations thereof.

The present invention further provides a method for searching nucleic acid databases comprising providing a central node comprising a processor, a plurality of sub-nodes in electronic communication with the central node, said sub-nodes comprising sequence database information, and nucleic acid sequence to be searched; providing the nucleic acid sequence to be searched to the central node; and concurrently sending the nucleic acid sequence information to be searched from the central node to the plurality of sub-nodes; and searching the sequence database information with the nucleic acid sequence to be searched to generate search results. In some embodiments, the method further comprises the step of sending the search results from the plurality of sub-nodes to the central node. In preferred embodiments, the latter steps are

complete in two seconds or less. In some embodiments, two or more distinct sequence databases are stored on the plurality of sub-nodes. In some embodiments, one of the two or more distinct sequence databases is stored on two or more of the plurality of sub-nodes. In some embodiments, two or more copies of the two or more distinct sequence databases are stored on the plurality of sub-nodes. In some embodiments, each of the plurality of sub-nodes comprises a single sequence database. In some embodiments, the nucleic acid sequence to be searched comprises a single nucleotide polymorphism or RNA. In some preferred embodiments, the sequence and variation in that sequence information comprises one or more databases comprising GoldenPath, GenBank, dbSNP, UniGene, LocusLink, The SNP Consortium, the Japanese SNP, and HGBASE SNP, Ensemble databases.

The present invention also provides a system or method used in one or more components hereof for characterizing a target sequence comprising: screening the target sequence for the presence of repeat sequences and heterologous sequences to generate a masked target sequence; searching a plurality of sequence databases with the masked target sequence to generate search result data; and generating a report comprising the search result data. In some embodiments, the plurality of sequence databases comprises one or more databases including, but not limited to, polymorphism databases, genome databases, linkage databases, and disease association databases (e.g., GoldenPath, GenBank, dbSNP, UniGene, LocusLink, and SNP Consortium databases). In some embodiments, the target sequence comprises a single nucleotide polymorphism. In some preferred embodiments, the report provides a reliability score, said reliability score representing a likelihood of success of detecting the target sequence performance in a detection assay. In some embodiments, the report indicates the presence or absence of the target sequence in one or more of the plurality of sequence databases. In some embodiments, the report indicates a position of the target sequence in a genome. In some embodiments, the report provides polymorphism information related to the target sequence.

The present invention further provides a database (e.g. used in one or more components hereof) comprising allele frequency information, said allele frequency information generated by a method comprising: producing a detection assay for detecting a target sequence; testing five or more target sequences from different subjects with the detection assay to produce assay data; and storing the assay data in a database, wherein the assay data is correlated to at least one

characteristic of the subjects. In some embodiments, the target sequence comprises a single nucleotide polymorphism. In some embodiments, the at least one characteristic of the subjects comprises subject age, sex, race or disease state.

The present invention also provides a method for collecting genomic information comprising, providing: a detection assay that detects the presence of a target nucleic acid sequence in a sample, a software application on a computer system of a user, said software application configured to receive detection assay data, a database on a computer system of a service provider, a communications network, and one or more samples comprising nucleic acid; treating the one or more samples with the detection assay to generate assay data; collecting the assay data with the software application; transmitting the assay data from the computer system of the user to the computer system of the service provider using the communications network; and storing the assay data in the database. In some embodiments, the target nucleic acid sequence comprises a single nucleotide polymorphism, wherein the detection assay detects the presence or absence of the single nucleotide polymorphism. The present invention also provides databases generated by such methods. The databases are used in one or more components hereof.

The present invention provides methods, systems, processes, and routines for developing and optimizing nucleic acid detection assays for use in basic research, clinical research, and for the development of clinical detection assays.

In some embodiments, the present invention provides methods comprising; a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and b) processing the target sequence information such that a primer set is generated, wherein the primer set comprises a forward and a reverse primer sequence for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set. It is also appreciated that, in one variant, a customer provided sequence, is automatically augmented

upstream and downstream to allow appropriate primer design using the methods and systems described herein.

In other embodiments, the present invention provides methods comprising; a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and b) processing the target sequence information such that a primer set is generated, wherein the primer set comprises a forward and a reverse primer sequence for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide G or T, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In particular embodiments, a method (including computer programs and routines that provide the following functionality) comprising; a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and b) processing the target sequence information such that a primer set is generated, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the 5' region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the 3' region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In other embodiments, the present invention provides methods (including routines that provide the following functionality) comprising a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and b) processing the target sequence information such that a

primer set is generated, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the 5' region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the 3' region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide G or T, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In particular embodiments, the present invention provides methods (and routines providing the following functionality) comprising a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises a single nucleotide polymorphism, b)determining where on each of the target sequences one or more assay probes would hybridize in order to detect the single nucleotide polymorphism such that a footprint region is located on each of the target sequences, and c) processing the target sequence information such that a primer set is generated, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In some embodiments, the present invention provides methods (and routines providing the following functionality) comprising a) providing target sequence information for at least Y target sequences, wherein each of the target sequences comprises a single nucleotide polymorphism, b) determining where on each of the target sequences one or more assay probes would hybridize in order to detect the single nucleotide polymorphism such that a footprint region is located on each of the target sequences, and c) processing the target sequence information such that a primer set is generated, wherein the primer set comprises; i) a forward

primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide T or G, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In certain embodiments, the primer set is configured for performing a multiplex PCR reaction that amplifies at least Y amplicons, wherein each of the amplicons is defined by the position of the forward and reverse primers. In other embodiments, the primer set is generated as digital or printed sequence information. In some embodiments, the primer set is generated as physical primer oligonucleotides. Using the methods, routines and components herein is it possible to generate 100-plex and greater PCR primer reactions.

In certain embodiments, N[3]-N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[3]-N[2]-N[1]-3' of any of the forward and reverse primers in the primer set. In other embodiments, the processing comprises initially selecting N[1] for each of the forward primers as the most 3' A or C in the 5' region. In certain embodiments, the processing comprises initially selecting N[1] for each of the forward primers as the most 3' G or T in the 5' region. In some embodiments, the processing comprises initially selecting N[1] for each of the forward primers as the most 3' A or C in the 5' region, and wherein the processing further comprises changing the N[1] to the next most 3' A or C in the 5' region for the forward primer sequences that fail the requirement that each of the forward primer's N[2]-N[1]-3' is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In other embodiments, the processing (preferably electronic) comprises initially selecting N[1] for each of the reverse primers as the most 3' A or C in the complement of the 3' region. In some embodiments, the processing comprises initially selecting N[1] for each of the reverse primers as the most 3' G or T in the complement of the 3' region. In further embodiments, the processing comprises initially selecting N[1] for each of the reverse primers as the most 3' A or C in the 3' region, and wherein the processing further comprises changing the N[1] to the next

most 3' A or C in the 3' region for the reverse primer sequences that fail the requirement that each of the reverse primer's N[2]-N[1]-3' is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In particular embodiments, the footprint region comprises a single nucleotide polymorphism. In some embodiments, the footprint comprises a mutation. In some embodiments, the footprint region for each of the target sequences comprises a portion of the target sequence that hybridizes to one or more assay probes configured to detect the single nucleotide polymorphism. In certain embodiments, the footprint is this region where the probes hybridize. In other embodiments, the footprint further includes additional nucleotides on either end.

In some embodiments, the processing (electronic in one variant of the invention) further comprises selecting N[5]-N[4]-N[3]-N[2]-N[1]-3' for each of the forward and reverse primers such that less than 80 percent homology with a assay component sequence is present. In preferred embodiments, the assay component is a FRET probe sequence. In certain embodiments, the target sequence is about 300-500 base pairs in length, or about 200-600 base pair in length. In certain embodiments, Y is an integer between 2 and 500, or between 2-10,000.

In certain embodiments, the processing (electronic in one variant of the invention) comprises selecting x for each of the forward and reverse primers such that each of the forward and reverse primers has a melting temperature with respect to the target sequence of approximately 50 degrees Celsius (e.g. 50 degrees, Celsius, or at least 50 degrees Celsius, and no more than 55 degrees Celsius). In preferred embodiments, the melting temperature of a primer (when hybridized to the target sequence) is at least 50 degrees Celsius, but at least 10 degrees different than a selected detection assay's optimal reaction temperature.

In some embodiments, the forward and reverse primer pair optimized concentrations are determined for the primer set. In other embodiments, the processing is automated. In further embodiments, the processing is automated with a processor.

In other embodiments, the present invention provides a kit comprising the primer set generated by the methods of the present invention, and at least one other component (e.g. cleavage agent, polymerase, INVADER oligonucleotide, or other detection assay or detection assay component in another variant of the invention). In certain embodiments, the present

invention provides compositions comprising the primers and primer sets generated by the methods of the present invention.

In particular embodiments, the present invention provides methods (and routines utilizing methodology) comprising; a) providing; i) a user interface configured to receive sequence data, ii) a computer system having stored therein a multiplex PCR primer software application, and b) transmitting the sequence data from the user interface to the computer system, wherein the sequence data comprises target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and c) processing the target sequence information with the multiplex PCR primer pair software application to generate a primer set, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In some embodiments, the present invention provides methods (and routines used in the methodology) comprising; a) providing; i) a user interface configured to receive sequence data, ii) a computer system having stored therein a multiplex PCR primer software application, and b) transmitting the sequence data from the user interface to the computer system, wherein the sequence data comprises target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, and c) processing the target sequence information with the multiplex PCR primer pair software application to generate a primer set, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a

portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide G or T, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set.

In certain embodiments, the present invention provides systems comprising; a) a computer system (and routines used in the methodology) configured to receive data from a user interface, wherein the user interface is configured to receive sequence data, wherein the sequence data comprises target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region immediately downstream of the footprint region, b) a multiplex PCR primer pair software application operably linked to the user interface, wherein the multiplex PCR primer software application is configured to process the target sequence information to generate a primer set, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set, and c) a computer system having stored therein the multiplex PCR primer pair software application, wherein the computer system comprises computer memory and a computer processor.

In other embodiments, the present invention provides systems comprising; a) a computer system or computer configured to receive data from a user interface, wherein the user interface is configured to receive sequence data, wherein the sequence data comprises target sequence information for at least Y target sequences, wherein each of the target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of the footprint region, and iii) a 3' region

immediately downstream of the footprint region, b) a multiplex PCR primer pair software application operably linked to the user interface, wherein the multiplex PCR primer software application is configured to process the target sequence information to generate a primer set, wherein the primer set comprises; i) a forward primer sequence identical to at least a portion of the target sequence immediately 5' of the footprint region for each of the Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of the target sequence immediately 3' of the footprint region for each of the at least Y target sequences, wherein each of the forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide G or T, and N[2]-N[1]-3' of each of the forward and reverse primers is not complementary to N[2]-N[1]-3' of any of the forward and reverse primers in the primer set, and c) a computer system having stored therein the multiplex PCR primer pair software application, wherein the computer system comprises computer memory and a computer processor. In certain embodiments, the computer system is configured to return the primer set to the user interface.

The present invention relates to novel methods of producing oligonucleotides. In particular, the present invention provides an efficient, safe, and automated process for the production of large quantities of oligonucleotides.

In some embodiments, the present invention provides high-throughput oligonucleotide production systems comprising: an oligonucleotide synthesizer component, wherein the oligonucleotide synthesizer component comprises at least 100 oligonucleotide synthesizers. In particular embodiments, the system further comprises at least one oligonucleotide processing component. In certain embodiments, the system further comprises a centralized control network operably linked to the oligonucleotide synthesizer component.

In particular embodiments, the present invention provides methods for the high through-put production of oligonucleotides comprising; a) providing an oligonucleotide synthesizer component; and b) generating a high through-put quantity of oligonucleotides with the oligonucleotide synthesizer component, wherein the high through-put quantity comprises at least 1 per hour (e.g. at least 1, 10, 100, 1000, etc, per hour).

In some embodiments, the present invention provides methods for the production of an oligonucleotide comprising: a) providing; i) a first computer memory device comprising oligonucleotide specification information, and ii) an oligonucleotide synthesizer component, wherein the oligonucleotide synthesizer component comprises a) at least 100 oligonucleotide synthesizers (in another variant the number of synthesizers can be in the range of about 20 to about 1000 synthesizers depending on the number of syntheses each synthesizer is capable of executing), and b) a second computer memory device; and b) conveying the oligonucleotide specification information from the first computer memory device to the second computer memory device under conditions such that the oligonucleotide synthesizer component generates at least one oligonucleotide (e.g. at least 1, 10, 100, 1000, etc). In another variant of the invention where high throughput synthesizers are used it is possible to substitute fewer synthesizers but still accomplish a desired level of syntheses.

In certain embodiments, the present invention provides oligonucleotide production systems comprising: a) an oligonucleotide production component configured for divergent production of a set of oligonucleotides, wherein the set of oligonucleotides comprises first and second corresponding oligonucleotides, and wherein the oligonucleotide production component comprises first and second oligonucleotide manufacturing components; and b) a centralized control network operably linked to the oligonucleotide production component, wherein the centralized control network is configured for controlling the divergent production of the set of oligonucleotides.

In other embodiments, the present invention provides methods for the divergent production of oligonucleotides comprising; a) providing an oligonucleotide production component comprising an oligonucleotide synthesizer component and at least one oligonucleotide processing component; and b) employing the oligonucleotide production component for divergent production of a set of oligonucleotides, wherein the set of oligonucleotides comprises first and second corresponding oligonucleotides.

In some embodiments, the present invention provides high-throughput oligonucleotide purification systems comprising a plurality of HPLC devices operably connected to a single sample injector. In other embodiments, the system further comprises a centralized control network.

In particular embodiments, the present invention provides methods for the high-throughput purification of oligonucleotides comprising: a) providing; i) an oligonucleotide purification component comprising a plurality of HPLC devices operably connected to a single sample injector, and ii) an oligonucleotide sample comprising full-length oligonucleotides and truncated oligonucleotides; and b) processing the sample with the oligonucleotide purification component under conditions such that at least a portion of the truncated oligonucleotides are removed from the oligonucleotide sample.

In some embodiments, the present invention provides high-throughput oligonucleotide production systems comprising; a) an oligonucleotide production component comprising first and second oligonucleotide manufacturing components; and b) a sample rack configured for use in the first and second oligonucleotide manufacturing components without modification. In particular embodiments, the system further comprises a central reagent supply network.

In certain embodiments, the present invention provides methods for high-throughput processing of oligonucleotide samples, comprising: a) providing; i) an oligonucleotide production component comprising first and second manufacturing components, and ii) a sample rack integrated with the first manufacturing component, wherein the sample rack is configured for use in the first and second oligonucleotide manufacturing components without modification, and wherein the sample rack comprises a plurality of oligonucleotide samples; and b) processing at least a portion of the plurality of oligonucleotide samples with the first manufacturing component, c) transferring the sample rack from the first manufacturing component to the second manufacturing component; and d) processing at least a portion of the oligonucleotide samples with the second manufacturing component.

In particular embodiments, the present invention provides high-throughput oligonucleotide dry-down systems comprising a centrifugal evaporator configured for processing at least 1 aqueous oligonucleotide sample in one hour or less. In particular embodiments, the system is configured for processing at least 5 oligonucleotide samples per hour (e.g. 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, or more than 50). In different embodiments, the present invention provides high-throughput oligonucleotide dry down systems comprising a centrifugal evaporator configured for processing a plurality of oligonucleotide samples in one hour or less, wherein the

plurality of oligonucleotide samples comprises at least 1 liter of water (e.g. 1, 5, 10, 15, 35 or 50 liters of water).

In some embodiments, the present invention provides methods for the high-throughput dry-down of oligonucleotides comprising: a) providing; i) an oligonucleotide dry-down component comprising a centrifugal evaporator, and ii) a plurality of oligonucleotide samples comprising at least 10 aqueous oligonucleotide samples; and b) processing the plurality of oligonucleotide samples with the oligonucleotide dry-down component, wherein the processing renders each of the aqueous oligonucleotide samples substantially water-free in one hour or less.

In certain embodiments, the present invention provides methods for the high-throughput dry-down of oligonucleotides comprising: a) providing; i) an oligonucleotide dry-down component comprising a centrifugal evaporator, and ii) a plurality of aqueous oligonucleotide samples, wherein the plurality of oligonucleotide samples comprises at least one liter of water, and b) processing the plurality of oligonucleotide samples with the oligonucleotide dry-down component, wherein the processing renders the plurality of aqueous oligonucleotide samples substantially water-free in one hour or less.

In some embodiments, the present invention provides high-throughput oligonucleotide de-salting systems comprising an oligonucleotide de-salting component configured for processing at least 150 oligonucleotide samples per half hour. In particular embodiments, the oligonucleotide de-salting component comprises a robotic oligonucleotide sample handling device, and a sample rack.

In other embodiments, the present invention provides methods for the high-throughput de-salting of oligonucleotides comprising: a) providing; i) an oligonucleotide de-salting component comprising a robotic oligonucleotide sample handling device, and ii) a plurality of oligonucleotide samples comprising at least 150 oligonucleotide samples; and b) processing the plurality of oligonucleotide samples with the oligonucleotide de-salting component, wherein the processing renders each of the oligonucleotide samples substantially salt-free in a half-hour or less.

In other embodiments, the present invention provides high-throughput oligonucleotide dilute and fill systems comprising an oligonucleotide dilute and fill component, wherein the

oligonucleotide dilute and fill component comprises an automated liquid processing device operably linked to a spectrophotometer.

In some embodiments, the present invention provides methods method for the high-throughput dilute and fill of oligonucleotide samples comprising: a) providing; i) an oligonucleotide dilute and fill component comprising an automated liquid processing device operably linked to a spectrophotometer, and ii) a plurality of oligonucleotide samples; and b) processing the plurality of oligonucleotide samples with the oligonucleotide dilute and fill component, wherein the processing normalizes each of the oligonucleotide samples. It is appreciated that normalization of concentration is an important aspect of the invention with respect to the production of detection assays. In one variant, oligonucleotide production samples have their concentrations normalized. This normalization can be accomplished via the utilization of known extinction coefficient methods and knowledge of the sequence from production information.

The present invention also provides a nucleic acid synthesis reagent delivery system comprising: one or more reagent containers containing nucleic acid synthesis reagent; a branched delivery component attached to said one or more reagent containers such that the nucleic acid synthesis reagent can pass from said reagent containers to said branched delivery component, wherein the branched delivery component comprises a plurality of branches; and a plurality of delivery lines, the plurality of delivery lines attached on one end to a branch of the branched delivery component and attached on a second end to a nucleic acid synthesizer. The present invention is not limited by the number branches or delivery lines. In some embodiments, the plurality of branches comprises ten or more branches. In some embodiments, the plurality of delivery lines comprises ten or more delivery lines. In some embodiments, the branched delivery component comprises a sight glass. In some preferred embodiments, the sight glass comprises a purge valve. In yet other embodiments, the one or more of the plurality of delivery lines comprises a shut-off valve.

The present invention further provides a waste disposal system comprising: a waste tank comprising a waste input channel configured to receive liquid waste product and a waste output channel configured to remove liquid waste when the waste tank is purged; and a pressurized gas line attached to the waste tank, the pressurized gas line configured to deliver gas into the waste

tank when the waste tank is to be purged, wherein the gas line is configured to deliver a gas that allows purging of the waste tank. In some embodiments, the pressurized gas line is attached to an argon gas source. In preferred embodiments, the gas is delivered at a low pressure (e.g., 3-10 pounds per square inch). In some embodiments, the waste input channel is attached to a waste line, wherein the waste line is attached to a plurality of nucleic acid synthesizers (e.g., 20 or more nucleic acid synthesizers). In some preferred embodiments, the waste tank comprises a sight glass. In other preferred embodiments, the system further comprises an automated purge component, said automated purge component capable of detecting waste levels in the waste tank and purging the waste tank when the waste levels are at or above a threshold level (e.g., a pre-selected threshold level).

The present invention also provides a method for purifying nucleic acids comprising providing: an nucleic acid purification column, a buffer, and a nucleic acid mixture; contacting the nucleic acid mixture with the nucleic acid purification column; and adding the buffer to the nucleic acid purification column, wherein a nucleic acid molecule having between 23-39 nucleotides is eluted from the nucleic acid purification column in less than forty minutes, and in one variant of the invention can be accomplished in less than about 25 minutes. In some embodiments, the nucleic acid purification column is contained in an HPLC apparatus.

The present invention further provides a method for deprotecting nucleic acid molecules comprising providing: a multiwell plate configured to hold a plurality of protected nucleic acid molecules and a plurality of different protected nucleic acid molecules; placing the nucleic acid molecules into the multiwell plates; and treating the plate under conditions that resulted in the deprotection of the nucleic acid molecules. In some embodiments, the multiwell plate comprises a 96-well plate.

The present invention relates to nucleic acid synthesizers and methods of using and modifying nucleic acid synthesizers. For example, the present invention provides highly efficient, reliable, and safe synthesizers that find use, for example, in high throughput and automated nucleic acid synthesis, as well as methods of modifying pre-existing synthesizers to improve efficiency, reliability, and safety. The present invention also relates to synthesizer arrays for efficient, safe, and automated processes for the production of large quantities of oligonucleotides.

In some embodiments, the present invention provides systems comprising a synthesis and purge component, the synthesis and purge component comprising a cartridge and a drain plate, wherein the cartridge is configured to hold one or more nucleic acid synthesis columns and wherein the cartridge is separated from the drain plate by a drain plate gasket. In certain embodiments, the cartridge is configured to hold a plurality of nucleic acid synthesis columns. In particular embodiments, the cartridge is configured to hold 12 or more nucleic acid synthesis columns. In other embodiments, the cartridge is configured to hold 48 or more nucleic acid synthesis columns. In additional embodiments, the cartridge is configured to hold exactly 48 nucleic acid synthesis columns.

In some embodiments, the assembly comprising the cartridge, the drain plate and the drain plate gasket is configured to provide a substantially airtight seal between the assembly and the outside of each nucleic acid synthesis column. In one embodiment, the airtight seal between the assembly and each column is provided by an O-ring. In a preferred embodiment, each O-ring is positioned between the cartridge and the exterior surface of a column. In yet another variant, any material that provides a compressible interface can be used in the invention.

In certain embodiments, the drain plate gasket provides a substantially airtight seal between the cartridge and the drain plate. In other embodiments, the drain plate gasket provides an airtight seal between the cartridge and the drain plate. In some embodiments, the drain plate gasket comprises one or more alignment markers configured to allow aligned attachment of said cartridge to said drain plate. In additional embodiments, the drain plate gasket comprises one or more alignment markers configured to allow aligned attachment of the drain plate gasket to the cartridge. In other embodiments, the drain plate gasket comprises one or more alignment markers configured to allow aligned attachment of the gasket to the drain plate. In certain embodiments, the drain plate gasket comprises at least one drain cut-out. In other embodiments, the drain plate gasket comprises at least four drain cut-outs. In still other embodiments, the drain plate gasket comprises one drain cut out for every synthesis column in the cartridge. In yet other embodiments, the cut outs in the drain plate gasket for each synthesis column are configured to provide an airtight seal between the outside of each nucleic acid synthesis column and the assembly comprising the cartridge, the drain plate, and the drain plate gasket.

In some embodiments, the present invention provides systems comprising a synthesis and purge component, the synthesis and purge component comprising a cartridge and a drain plate, wherein the cartridge is configured to hold one or more nucleic acid synthesis columns and wherein the cartridge is separated from the drain plate by a drain plate gasket. In some embodiments, the drain plate comprises at least one drain (*e.g.* 1, 2, 3, 4, 5, 10,... 20, ...). In other embodiments, the system further comprises a waste tube, the waste tube comprising input and output ends, wherein the input end is configured to receive waste materials from the drain. In particular embodiments, the waste tube comprises an inner diameter of at least 0.187 inches (preferably at least 0.25 inches). In some embodiments, the waste tube and the drain are configured such that, when the drain is contacted with the waste tube for waste removal, the waste tube encloses at least a portion of the drain (*See, e.g.*, Figure 40). In particular embodiments, the drain forms a sealed contact point with an interior portion of the waste tube when the drain is enclosed in the waste tube. In still other embodiments, the drain further comprises a drain sealing ring. In certain embodiments, the system further comprises a waste valve wherein the waste valve is configured to receive waste from the output end of the waste tube. In particular embodiments, the waste valve comprises an interior diameter of at least 0.187 inches (preferably at least 0.25 inches). In some embodiments, the waste valve provides a straight-through path for the waste (*e.g.* as opposed to an angled path). Straight-through paths can be accomplished, for example, by the use of a gate or ball valve.

In some embodiments, the system further comprises a plurality of dispense lines, the dispense line configured for delivering at least one reagent to a synthesis column in the cartridge. In certain embodiments, the dispense lines comprise an interior diameter of at least 0.25 mm. In particular embodiments, the system further comprises an alignment detector. In particular embodiments, the alignment detector is configured to detect the alignment of a waste tube and a drain. In other embodiments, the alignment detector is configured to detect the alignment of a dispense line and a receiving hole of the cartridge. In some embodiments, the alignment detector is configured to detect a tilt alignment of the synthesis and purge component.

In some embodiments, the system of the present invention further comprises a motor attached to the synthesis and purge component and configured to rotate the synthesis and purge component. In particular embodiments, the motor is attached to the synthesis and purge

component by a motor connector. In further embodiments, the system further comprises a bottom chamber seal positioned between the motor connector and the synthesis and purge component. In certain embodiments, the system of the present invention comprises two drain. In preferred embodiments, the two drain are located on opposite sides of the drain plate.

In some embodiments of the systems of the present invention, the synthesis and purge component is contained in a chamber. In certain embodiments, a chamber bowl and a top cover (when in place) combine to form a chamber (*e.g.* which may be pressurized, for example, with inert gas). One example is depicted in Figure 34 where chamber bowl 18 and top cover 30 combine to form an exemplary chamber. In some embodiments, the chamber comprises a bottom surface (*e.g.* bottom of a chamber bowl, see, *e.g.* Figure 41) comprising the top portion of two waste tubes (which may, for example, extend downward from bottom of the chamber). In preferred embodiments, the waste tubes are positioned symmetrically on the bottom surface of the chamber (*see, e.g.*, Figure 41).

In particular embodiments, the systems of the present invention further comprise a chamber drain having open and closed positions, the chamber drain configured to allow gas emissions (or liquid waste) to pass out of the chamber when in the open position.

In some embodiments, the systems of the present invention further comprise a reagent dispensing station, wherein the reagent dispensing station is configured to house one or more reagent reservoirs, such that reagents in reagent reservoirs can be delivered to the cartridge. In certain embodiments, the reagent dispensing station comprises one or more ventilation tubes (*e.g.*, connected to one or more ventilation valves of the reagent dispensing station) configured to remove gaseous emissions from the reagent dispensing station. In certain embodiments, the reagent dispensing station provides an enclosure. In preferred embodiments, the enclosure comprises a viewing window to allow visual inspection of the reagent reservoirs without opening the enclosure. In preferred embodiments, one reagent dispensing station is configured to serve multiple synthesizers.

In particular embodiments, the systems of the present invention are capable of maintaining a gas pressure in the chamber sufficient to purge synthesis columns prior to addition of reagents to the synthesis columns.

In some embodiments, the nucleic acid synthesis systems of the present invention comprise a cartridge in a chamber, the cartridge comprising a plurality of synthesis columns, wherein the synthesis columns contain packing material that provides a resistance against pressurized gas contained in the chamber, the resistance being sufficient to maintain a pressure in the chamber that is capable of purging synthesis columns prior to addition of reagents to the synthesis columns. In certain embodiments, one or more of the plurality of synthesis columns does not undergo a synthesis reaction. In particular embodiments, two or more different lengths of oligonucleotides are synthesized in the plurality of synthesis columns. In other embodiments, the packing material comprises a frit. In some embodiments, the frit is a bottom frit. In other embodiments, the frit is a top frit. In preferred embodiments, the packing material comprises a top frit, solid support, and a bottom frit. In particularly preferred embodiments, the solid support is polystyrene. In some embodiments, the packing material comprises a synthesis matrix.

In some embodiments, the present invention provides nucleic acid synthesis systems comprising a synthesis and purge component in a pressurized chamber, the synthesis and purge component comprising a plurality of synthesis columns, wherein the synthesis columns contain packing material sufficient to maintain pressure in the chamber during a purging operation to purge liquid reagent from the plurality of synthesis columns when at least one of the plurality of synthesis columns does not contain liquid reagent. In certain embodiments, more than one of the plurality of synthesis columns (*e.g.* 2, 3, 5, 10) do not contain liquid reagent (and the remaining synthesis columns do contain liquid reagent).

In certain embodiments, the present invention provides nucleic acid synthesis systems comprising: a) a synthesis and purge component, the synthesis and purge component comprising a cartridge and a drain plate separated by a drain plate gasket, wherein the cartridge is configured to hold twelve or more nucleic acid synthesis columns; b) a drain positioned in the drain plate; c) a chamber comprising an inner surface, the chamber housing the synthesis and purge component and the drain; d) a waste tube, the waste tube comprising input and output ends, wherein the input end is configured to receive waste materials from the drain, wherein the waste tube comprises an inner diameter of at least 0.187 inches; e) a waste valve configured to receive waste from the output end of the waste tube, wherein the waste valve comprises in interior diameter of at least 0.187 inches; f) a reagent dispensing station, wherein the reagent dispensing station is

configured to house one or more reagent reservoirs; g) a plurality of dispense lines, the dispense lines configured for delivering reagents from the reagent reservoirs to a synthesis column in the cartridge, wherein the dispense lines comprise an interior diameter of at least 0.25 mm) a rotating motor attached to the synthesis and purge component by a motor connector and configured to rotate the synthesis and purge component; and i) a gas line configured to release gas into the chamber to create a gas pressure in the chamber greater than a gas pressure in the waste tube. In certain embodiments, the system is capable of maintaining gas pressure in the chamber at a sufficient level to purge the synthesis columns prior to addition of reagents to the synthesis columns.

In some embodiments, the synthesizer further comprises providing energy, such as heat, to the synthesis columns. Heating of the synthesis column finds use, for example, in decreasing the coupling time during a nucleic acid synthesis. It can also broaden the range of the chemical protocols that can be used in high throughput synthesis, *e.g.* by improving the efficiency of less efficient chemistries, such as the phosphate triester method of oligonucleotide synthesis. In other embodiments, the synthesizer further comprises a mixing component, such as an agitator, configured to agitate the synthesis columns (*e.g.*, to mix reaction components, and to facilitate mass exchange between the reaction medium and the solid support).

In some embodiments, the present invention provides methods for synthesizing nucleic acids comprising: a) providing: i) a nucleic acid synthesizer comprising a synthesis and purge component, the synthesis and purge component comprising a cartridge and a drain plate, wherein the cartridge holds a plurality of nucleic acid synthesis columns and wherein the cartridge is separated by a drain plate gasket from the drain plate, and ii) nucleic acid synthesis reagents; and b) introducing a portion of the nucleic acid synthesis reagents into at least one of the nucleic acid synthesis columns to provide a first synthesis reaction; c) purging the nucleic acid synthesis columns by creating a pressure differential across the nucleic acid synthesis columns; and d) introducing a second portion of the nucleic acid synthesis reagents into at least one of the nucleic acid synthesis columns to provide a second synthesis reaction. In particular embodiments, the drain plate gasket provides a substantially airtight seal between the cartridge and the drain plate. In other embodiments, the drain plate gasket provides an airtight seal between the cartridge and the drain plate.

The present invention further provides a cartridge for use in an open nucleic acid synthesis system, said cartridge comprising a plurality of receiving holes configured to hold nucleic acid synthesis columns, wherein the cartridge is further configured to receive one or more O-rings, wherein the presence of the one or more O-rings provides a seal between the nucleic acid synthesis columns and the plurality of receiving holes (*i.e.*, the O-ring contacts an interior wall of the receiving hole and an exterior wall of the synthesis column to form a seal). In some embodiments, the cartridge is provided as part of a nucleic acid synthesis system. The present invention is not limited by the nature of the O-ring. For example, in some embodiments, the cartridge is associated with a gasket, wherein the gasket provides the O-rings (*e.g.*, through one or more holes in the gaskets, such that when the gasket is associated with the cartridge [*e.g.*, affixed to an outer surface of the cartridge] a seal is formed between the a receiving hole of the cartridge and a synthesis column within the receiving hole [*see e.g.*, Figure 46C]). In other embodiments, the O-ring is provided in a groove within the receiving hole. For example, in some embodiments, the groove is located at the top surface of the receiving hole. In such embodiments, the plurality of receiving holes comprise an upper portion and a lower portion, wherein the lower portion comprises a first diameter and the upper portion comprises a second diameter that is larger than the first diameter (*see e.g.*, Figure 46A). In other embodiments, the groove is located within an interior portion of the receiving hole. In such embodiments, the plurality of receiving holes comprise an upper portion with a first diameter, a middle portion with a second diameter, and a lower portion with a third diameter, wherein the second diameter is larger than the first diameter and larger than the third diameter (the first and third diameters may be the same as each other or different). When an O-ring is placed in the groove, the O-ring contains an internal diameter less than the first diameter and less than the third diameter, such that it can contact a synthesis column placed within the receiving hole (*see e.g.*, Figure 46B).

In some embodiments, the cartridge comprises a rotary cartridge. In some preferred embodiments, O-rings are provided in the cartridge. In some preferred embodiments, the O-ring is configured to form a substantially airtight or pressure-tight seal between the receiving hole and the nucleic acid synthesis column, when said nucleic acid synthesis column is present.

The present invention further provides a nucleic acid synthesis system comprising a synthesis and purge component in a pressurizable chamber, said synthesis and purge component

comprising a cartridge, wherein the cartridge in configured to hold a plurality of nucleic acid synthesis columns, and wherein said cartridge is further configured to provides seals between said cartridge and each of said plurality of nucleic acid synthesis columns so as to maintain pressure in said chamber during a purging operation to purge liquid reagent from said plurality of synthesis columns. In some embodiments, each of the seals between the cartridge and the plurality of nucleic acid synthesis columns is provided by an O-ring.

In some embodiments, the present invention provides a nucleic acid synthesizer comprising a plurality of synthesis columns and an energy input component that imparts energy to said plurality of synthesis columns to increase nucleic acid synthesis reaction rate in said plurality of synthesis columns. In some embodiments, said energy input component comprises a heating component. In preferred embodiments, said heating component provides substantially uniform heat. In some embodiments, said energy input component provides heated reagent solutions to said plurality of synthesis columns. In other embodiments, said energy input component comprises a heating coil. In yet other embodiments, said energy input component comprises a heat blanket. In yet other embodiments, said heating component comprises a resistance heater, a Peltier device, a magnetic induction device or a microwave device. In still other embodiments, said energy input component comprises a heated room. In further embodiments, said energy input component provides energy in the electromagnetic spectrum. In yet other embodiments, said energy input component comprises an oscillating member. In some embodiments, said energy input component provides a periodic energy input, and in other embodiments, said energy input component provides a constant energy input.

In some preferred embodiments, said energy input heats said plurality of synthesis columns in the range of about 20 to about 60 degrees Celsius.

In some embodiments, the present invention provides a nucleic acid synthesizer comprising a fail-safe reagent delivery component configured to deliver one or more reagent solutions to said plurality of synthesis columns. In some embodiments, the fail-safe reagent delivery component comprises a plurality of reagent tanks. In preferred embodiments, said plurality of reagent tanks comprise one or more tanks selected from the group consisting of acetonitrile tanks, phosphoramidite tanks, argon gas tanks, oxidizer tanks, tetrazole tanks, and capping solution tanks. In some particularly preferred embodiments, said reagent tanks comprise

a plurality of large volume containers, each said large volume container comprising at least one of said reagent solutions. In some embodiments, the present invention provides high-throughput oligonucleotide production systems comprising: an oligonucleotide synthesizer array, wherein the oligonucleotide synthesizer array comprises at least 5 oligonucleotide synthesizers. In preferred embodiments, the oligonucleotide synthesizer array comprises at least 10 or at least 100 oligonucleotide synthesizers. In certain embodiments, the system further comprises a centralized control network operably linked to the oligonucleotide synthesizer component.

In particular embodiments, the present invention provides methods for the high through-put production of oligonucleotides comprising; a) providing an oligonucleotide synthesizer array; and b) generating a high through-put quantity of oligonucleotides with the oligonucleotide synthesizer array, wherein the high through-put quantity comprises at least 1 per hour (*e.g.* at least 1, 10, 100, 1000, etc, per hour).

The present invention provides a production facility comprising an array of synthesizers. In some embodiments, the production facility of the present invention comprises a fail-safe reagent delivery system. In other embodiments, the production facility of the present invention comprises a centralized waste collection system. In yet other embodiments, the production facility of the present invention comprises a centralized control system. In preferred embodiments, the production facility of the present invention comprises a fail-safe reagent delivery system, a centralized waste collection system and a centralized control system.

In some embodiments, the present invention provides an automated production process. In some embodiments, the automated production process includes an oligonucleotide synthesizer component and an oligonucleotide-processing component.

The present invention also provides integrated systems that link nucleic acid synthesizers to other nucleic acid production components. For example, the present invention provides a system comprising a nucleic acid synthesizer and a cleavage and deprotect component. In some embodiments, the synthesizer is configured for parallel synthesis of nucleic acid molecules in three or more synthesis columns. In some embodiments, the system further comprises sample tracking software configured to associate sample identification tags (*e.g.*, electronic identification numbers, barcodes) with samples that are processed by the nucleic acid synthesizer and the cleavage and deprotect component. In some preferred embodiments, the sample tracking

software is further configured to receive synthesis request information from a user, prior to sample processing by the nucleic acid synthesizer. In some embodiments, the system further comprises a robotic component configured to transfer columns from the nucleic acid synthesizer to the cleavage and deprotect component. In other preferred embodiments, the robotic component is further configured to transfer the columns from the cleavage and deprotect component to a purification component and/or to additional production components described herein.

The present invention also provides control systems for operating one or more components of the systems of the present invention. For example, the present invention provides a system comprising a processor, wherein the processor is configured to operate a nucleic acid synthesizer for parallel synthesis of three or more nucleic acid molecules. The present invention further provides a system comprising a processor, wherein said processor is configured to operate a nucleic synthesizer and a cleavage and deprotect component. In some embodiments, the system further comprises a computer memory, wherein the computer memory comprises nucleic acid sample order information (e.g., information obtained from a user specifying the identity of a polymer to be synthesized and/or specifying one or more characteristics of the polymer such as sequence information). In some embodiments, the computer memory further comprises allele frequency information and/or disease association information.

In some embodiments, the present invention provides oligonucleotide synthesizers comprising a reaction chamber and a lid, wherein in an open position, the lid provides a substantially enclosed ventilated workspace. In certain embodiments, the present invention provides methods of protecting an operator of an oligonucleotide synthesizer comprising channeling ambient air away from an operator toward an interior space of the synthesizer (e.g. down through the top surface, or up through the top cover). In other embodiments, the present invention provides apparatuses comprising, in combination, an oligonucleotide synthesizer and a venting hood. In some embodiments, the apparatuses are for production of oligonucleotides, wherein the apparatus comprises a venting component configured to draw air away from a reaction chamber of the apparatus. In certain embodiments, the present invention provides systems comprises a plurality of oligonucleotide apparatuses (e.g. e.g. at least 100 synthesizers).

In particular embodiments, the present invention provides a polymer synthesizer comprising a ventilated workspace. In some embodiments, certain embodiments, the polymer synthesizer is a nucleic acid synthesizer. In certain embodiments, the synthesizer comprises a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, wherein the top enclosure is configured for attachment to a top cover of a synthesizer to form a primarily enclosed space over the top cover. In other embodiments, the synthesizer comprises a base, wherein the base comprises a primarily enclosed space and a ventilation opening.

In certain embodiments, the top plate is configured for attachment to a ventilation tube such that air in the primarily enclosed space may be drawn through the ventilation opening into the ventilation tube. In other embodiments, the top plate further comprises an outer window, and wherein the ventilation opening is formed in the outer window. In certain embodiments, the top enclosure further comprises at least four sides (e.g. 4 sides, 5 sides, etc.). In certain embodiments, the top cover further comprises a ventilation slot.

In certain embodiments, the present invention provides polymer synthesizer (e.g. nucleic acid synthesizer) comprising; a) a top cover with a ventilation slot, and b) a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, and wherein the top enclosure is attached to the top cover to form a primarily enclosed space above the top cover.

In certain embodiments, the present invention provides a lid enclosure comprising; a) a top cover with a ventilation slot, and b) a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, and wherein the top enclosure is attached to the top cover to form a primarily enclosed space over the top cover. In certain embodiments, the top plate is configured for attachment to a ventilation tube. In particular embodiments, the top plate is configured for attachment to a ventilation tube such that air in the primarily enclosed space may be drawn through the ventilation opening into the ventilation tube. In other embodiments, the top cover is configured to attach to a top surface of a nucleic acid synthesizer with a chamber bowl.

In some embodiments, the ventilation slot is configured such that air in the chamber bowl may drawn in through the ventilation slot and into the primarily enclosed space. In other embodiments, the top plate further comprises an outer window, and wherein the ventilation

opening is formed in the outer window. In certain embodiments, the top enclosure further comprises at least four sides.

In certain embodiments, the present invention provides a polymer synthesizer (e.g., nucleic acid synthesizer) comprising; a) a top surface of a nucleic acid synthesizer, b) a lid enclosure comprising; i) a top plate with a ventilation opening, and ii) a top cover with a ventilation slot; and wherein the lid enclosure is attached to the top surface. In some embodiments, the lid enclosure is attached to the top surface by at least one hinge such that the lid enclosure may be raised and lowered. In certain embodiments, the present invention provides systems comprises a plurality of the polymer synthesizers (e.g., at least 100 synthesizers).

In some embodiments, the present invention provides side panels configured to extend between at least one side of a top cover (or lid enclosure) and a top surface of a nucleic acid synthesizer such that a barrier to air is created on at least one side of the synthesizer when the top cover is extended upward from the top surface. In other embodiments, the present invention provides a panel (e.g. front panel or side panel) configured to extend at least part way between at least one side of a top cover (or lid enclosure) and a top surface of a nucleic acid synthesizer such that at least a partial barrier to air is created on at least one side of the synthesizer when the top cover is extended upward such that it is not in contact with the top surface.

In other embodiments, the present invention provides polymer synthesizers (e.g. nucleic acid synthesizers) summary comprising; a) a top surface of a nucleic acid synthesizer, b) a lid enclosure comprising; i) a top plate with a ventilation opening, ii) a top cover with a ventilation slot; and iii) at least one top enclosure side; and c) a panel; wherein the lid enclosure is attached to the top surface by at least one hinge such that the lid enclosure may be raised and lowered, and wherein the panel is configured to extend (at least part way) between the at least one top enclosure side and the top surface such that at least a partial barrier to air is created when the lid enclosure is extended upward from the top surface. In certain embodiments, the present invention provides systems comprising a plurality of the polymer synthesizers (e.g., at least 100 synthesizers).

In particular embodiments, the present invention provides systems comprising; a) a ventilation tube, and b) a lid enclosure comprising; a) a top cover with a ventilation slot, and b) a top enclosure comprising a top plate with a ventilation opening, wherein the top enclosure is

attached to the top cover to form a primarily enclosed space over the top cover. In some embodiments, the systems further comprise a vacuum source (e.g. centralized vacuum system).

In certain embodiments, the top plate is configured for attachment to the ventilation tube. In other embodiments, the ventilation tube is configured for attachment to the vacuum source. In particular embodiments, the system further comprises a synthesis and purge component, the synthesis and purge component comprising a cartridge and a drain plate separated by a drain plate gasket, wherein the cartridge is configured to hold a plurality of nucleic acid synthesis columns. In some embodiments, the systems further comprise a plurality of dispense lines, wherein the plurality of dispense lines are located in the primarily enclosed space.

In certain embodiments, the systems further comprise at least one side panel, wherein the at least one side panel is configured to extend between at least one side of the lid enclosure and a top surface of a nucleic acid synthesizer (e.g., such that a barrier to air is created on at least one side of the synthesizer when the top cover is extended upward from the top surface).

In some embodiments, the present invention provides systems comprising; a) a nucleic acid synthesizer comprising; i) a top surface, and ii) a top cover comprising a ventilation slot, wherein the top cover is attached to the top surface by at least one hinge such that the top surface may be raised and lowered; and b) a panel configured to extend at least part way between at least one side of the top cover and the top surface such that at least a partial barrier to air is created on at least one side of the nucleic acid synthesizer when the top cover is extended upward. In other embodiments, the panel is configured to fully extend between the at least one side of the top cover and the top surface such that a complete barrier to air is created on at least one side of the nucleic acid synthesizer when the top cover is extended upward. In some embodiments, the panel comprises a side panel or a front panel.

In certain embodiments, the system further comprises a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, and wherein the top enclosure is attached to the top cover to form a primarily enclosed space over the top cover. In other embodiments, the system further comprises a ventilation tube. In particular embodiments, the system further comprises a vacuum source. In other embodiments, the vacuum source comprises a centralized vacuum system. In particular embodiments, the top plate is configured for

attachment to the ventilation tube. In certain embodiments, the ventilation tube is configured for attachment to the vacuum source.

In some embodiments, the present invention provides methods comprising forming a ventilation opening in a top plate of a top enclosure such that the top plate is configured for attachment to a ventilation tube. In certain embodiments, the present invention provides methods comprising; a) providing; i) a top enclosure comprising a top plate, and ii) a ventilation tube; and b) forming a ventilation opening in the top plate, and c) attaching the ventilation tube to the top plate such that the ventilation tube forms a seal around the ventilation opening. In further embodiments, the methods further comprise step d) attaching a least one panel to the top enclosure.

In other embodiments, the present invention provides methods comprising; a) providing; i) a top cover of a nucleic acid synthesizer comprising a ventilation slot, wherein the top cover is configured to be attached to a top surface of a nucleic acid synthesizer such that the top surface may be raised and lowered; and ii) a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, and b) attaching the top enclosure to the top cover such that a primarily enclosed space is formed over the top cover. In other embodiments, the methods further comprise the step of attaching at least one panel to the top enclosure (or the top cover), wherein the at least one panel extends at least part way between at least one side of the top cover (or the top cover) and the top surface such that at least a partial barrier to air is created on at least one side of the synthesizer when the top cover is extended upward such that it is not in contact with the top surface.

In particular embodiments, the present invention provides methods comprising; a) providing; i) a nucleic acid synthesizer comprising; i) a top cover with a ventilation slot, and ii) a top enclosure, wherein the top enclosure comprises a top plate with a ventilation opening, wherein the top enclosure is attached to the top cover to form a primarily enclosed space above the top cover, and wherein the top plate is attached to a ventilation tube such that the ventilation tube forms a seal around the ventilation opening, and ii) a vacuum source attached to the ventilation tube, and b) activating the vacuum source such that air is drawn into the ventilation slot, through the primarily open space, and out through the ventilation opening into the ventilation tube.

In some embodiments, the present invention provides kits comprising; a) a top enclosure comprising a top plate with a ventilation opening, wherein the top enclosure is configured for attachment to a top cover of a synthesizer to form a primarily enclosed space over the top cover, and b) a printed material component, wherein the printed material component comprises written instruction for installing the top enclosure onto the top cover.

In other embodiments, the present invention provides kits comprising; a) a panel configured to extend at least part way between at least one side of a top cover (or lid enclosure) and a top surface of a nucleic acid synthesizer such that at least a partial barrier to air is created on at least one side of the synthesizer when the top cover is extended upward such that it is not in contact with the top surface, and b) a printed material component, wherein the printed material component comprises written instructions for installing the panel onto a top cover (or lid enclosure).

The present invention relates to polymer synthesizers and methods of using polymer synthesizers. For example, the present invention provides highly efficient, reliable, and safe synthesizers that find use, for example, in high throughput and automated nucleic acid synthesis. The present invention also relates to synthesizer arrays for efficient, safe, and automated processes for the production of large quantities of oligonucleotides.

For example, the present invention provides a system comprising a closed system solid phase synthesizer configured for parallel synthesis (*e.g.*, simultaneous side-by-side synthesis) of three or more polymers (*e.g.*, 3, 4, 5, 6, 7, . . ., 10, . . ., 48, . . ., 96, . . .). The present invention is not limited by the nature of the polymer. Polymers include, but are not limited to, nucleic acids and polypeptides. In some preferred embodiments, the nucleic acid polymers comprise DNA. In some particularly preferred embodiments, the DNA comprises an oligonucleotide.

The synthesizers of the present invention allow parallel synthesis of multiple polymers. Each of the synthesized polymers may be identical to one another (*e.g.*, in composition, sequence, length, etc.) or may be different than one another (*e.g.*, in composition, sequence, length, etc.). Thus, the synthesizers of the present invention may be configured to simultaneously produce three or more distinct polymers (*e.g.*, oligonucleotides).

Because the synthesizers of the present invention allow parallel processing of polymers, large numbers of polymers may be produced in a single synthesizer in a short period of time.

For example, the synthesizer may be configured to produce 100 or more polymers per day. In some embodiments, the synthesizer may be configured to produce 1000-2000 or more polymers per day. For example, synthesizers may be configured to produce 2000 or more oligonucleotide per day (*e.g.*, oligonucleotides containing 20-40 or more bases). In some preferred embodiments, the produced polymers (*e.g.*, 2000 or more produced polymers) are produced at a 1 μM synthesis scale. In some embodiments, the produced polymers are produced on a micro-scale, *e.g.*, less than 5 nmole synthesis scale. In some preferred embodiments, micro-scale synthesis is performed on a 0.1 to 1 nmole synthesis scale.

The present invention also provides a solid phase synthesizer comprising: a reaction support comprising three or more (*e.g.*, 3, 4, 5, 6, 7, . . ., 10, . . ., 48, . . ., 96, . . .) reaction chambers (*e.g.*, chambers that are isolated from one another, such that fluid does not pass from one chamber to another during synthesis); and a plurality of reagent dispensers configured to simultaneously form closed fluidic connections with each of the reaction chambers, wherein the reagent dispensers are each configured to deliver all reagents necessary for a polymer synthesis reaction. In some embodiments, the reaction chambers comprise synthesis columns. For example, the reaction support provides a fixed surface to support three or more synthesis columns. In some embodiments, the synthesis columns comprise nucleic acid synthesis columns (*e.g.*, columns designed for use with EXPEDITE nucleic acid synthesizers [Applied Biosystems, Foster City, CA], 3900 High-Throughput Columns for use with the 3900 DNA Synthesizer [Applied Biosystems], DNA synthesis columns from Biosearch Technologies, Novato, CA). In preferred embodiments, the reaction support is configured to contain and form a tight seal around multiple, different synthesis columns (*e.g.*, of different sizes or from different manufacturers), so as to allow any number of commercially available columns to be used with the synthesizer.

In some embodiments, the reagent dispensers are fluidicly connected to a plurality of reagent tanks (*e.g.*, through tubing). In preferred embodiments, reagent dispensers are constructed from any substantially inert materials including, but not limited to, stainless steel, glass, Teflon, and titanium. Tanks include, but are not limited to, acetonitrile tanks, phosphoramidite tanks, argon gas tanks, oxidizer tanks, tetrazole tanks, and capping solution tanks. In some embodiments, the tanks are contained within the synthesizer. In other embodiments, the tanks are contained on an outer surface of the synthesizer. In some preferred

embodiments, tanks are provided separately from the synthesizer (*e.g.*, in a different room, such as an explosion-proof room). For example, in some embodiments, the present invention provides large volume synthesis facilities containing multiple synthesizers, wherein two or more of the synthesizer are serviced by the same reagent tanks. In some such embodiments, "large volume containers" are used as reagent tanks. Individual large volume reagent tanks contain from about 200 liters to about 2500 liters of acetonitrile, from about 200 liters to about 2500 liters of deblocking solution; from about 2 liters to about 200 liters of amidite; from about 20 liters to about 200 liters of activator (*e.g.*, tetrazol); from about 20 liters to about 200 liters of capping reagents; or from about 20 liters to about 200 liters of oxidizer. Alternatively, a plurality of tanks containing a combined capacity as indicated above may be used. In some embodiments, the large volume reagent tanks are connected to a plurality of synthesizers through a large volume reagent delivery system, which allows large volumes of reagents to be delivered simultaneously to each of the synthesizers

Various useful reagents and coupling chemistries are described in U.S. Pat. 5,472,672 to Bennan, and U.S. Pat. No. 5,368,823 to McGraw *et al.* (both of which are herein incorporated by reference in their entireties). In addition to phosphoramidite chemistries, phosphate and phosphite triester methods, and H-phosphonate methods of oligonucleotide synthesis are contemplated.

In some embodiments, the reaction support comprises a fixed reaction support (e.g., a reaction support that does not move during operation). In some embodiments, the reaction support comprises a plurality of waste channels. In preferred embodiments, the waste channels in closed fluidic contact with each of the reaction chambers (See *e.g.*, Figure 53).

In some embodiments, the synthesizer further comprises providing energy, such as heat to the reaction chambers. Heating of the reaction chamber finds use, for example, in decreasing the coupling time during a nucleic acid synthesis. It can also broaden the range of the chemical protocols that can be used in high throughput synthesis, *e.g.* by improving the efficiency of less efficient chemistries, such as the phosphate triester method of oligonucleotide synthesis. In other embodiments, the synthesizer further comprises a mixing component, such as an agitator, configured to agitate the reaction chambers (*e.g.*, to mix reaction components, and to facilitate mass exchange between the reaction medium and the solid support).

The present invention further provides a solid phase synthesizer comprising: a fixed reaction support comprising three or more reaction chambers; and a plurality of reagent dispensers configured to simultaneously form closed fluidic connections with each of said reaction chambers.

The present invention also provides integrated systems that link nucleic acid synthesizers to other nucleic acid production components. For example, the present invention provides a system comprising a closed system nucleic acid synthesizer and a cleavage and deprotect component. In some embodiments, the synthesizer is configured for parallel synthesis of nucleic acid molecules at three or more reaction sites. In some preferred embodiments, the system further comprises a reaction support comprising three or more reaction chambers, wherein the reaction support is configured for operation with both the nucleic acid synthesizer and the cleavage and deprotect component. In some embodiments, the system further comprises sample tracking software configured to associate sample identification tags (*e.g.*, electronic identification numbers, barcodes) with samples that are processed by the nucleic acid synthesizer and the cleavage and deprotect component. In some preferred embodiments, the sample tracking software is further configured to receive synthesis request information from a user, prior to sample processing by the nucleic acid synthesizer. In some embodiments, the system further comprises a robotic component configured to transfer the reaction support from the nucleic acid synthesizer to the cleavage and deprotect component. In other preferred embodiments, the robotic component is further configured to transfer the reaction support from the cleavage and deprotect component to a purification component and/or to additional production components described herein.

The present invention also provides control systems for operating one or more components of the systems of the present invention. For example, the present invention provides a system comprising a processor, wherein the processor is configured to operate a close system nucleic acid synthesizer for parallel synthesis of three or more nucleic acid molecules. The present invention further provides a system comprising a processor, wherein said processor is configured to operate a nucleic synthesizer and a cleavage and deprotect component. In some embodiments, the system further comprises a computer memory, wherein the computer memory comprises nucleic acid sample order information (e.g., information obtained from a user

specifying the identity of a polymer to be synthesized and/or specifying one or more characteristics of the polymer such as sequence information). In some embodiments, the computer memory further comprises allele frequency information and/or disease association information.

In some embodiments, the present invention relates to detecting mutations in pooled nucleic acid samples. In particular, the present invention relates to compositions and methods for detecting mutations or measuring allele frequencies in pooled nucleic acid samples employing the INVADER detection assay or other detection assays described herein. In some embodiments, the present invention provides methods for detecting an allele frequency of a polymorphism, comprising: a) providing; i) a pooled sample, wherein the pooled sample comprises target nucleic acid sequences from at least 10 individuals (or at least 50, or at least 100, or at least 250, or at least 500, or at least 1000 individuals, etc.); and ii) INVADER detection reagents (e.g. primary probes, INVADER oligonucleotides, FRET cassettes, a structure specific enzyme, etc.) configured to detect the presence or absence of a polymorphism; and b) contacting the pooled sample with the INVADER detection reagents to generate a detectable signal; and c) measuring the detectable signal, thereby determining a number of the target nucleic acid sequences that contain the polymorphism (e.g. a quantitative number of molecules, or the allele frequency for the polymorphism in a population, is determined). In some embodiments, signals from two or more alleles for a particular target nucleic acid locus are measured and the numbers are compared. In preferred embodiments, the measurements for two or more different alleles of a particular target nucleic acid locus are measured in a single reaction. In other embodiments, measurements from one or more alleles of a particular target nucleic acid locus are compared to measurements from one or more reference target nucleic acid loci. In preferred embodiments, measurements from one or more alleles of a particular target nucleic acid locus are compared to measurements from one or more reference target nucleic acid loci in the same reaction mixture. Further methods allow a single individual's particular allele frequency (i.e., frequency of the mutation among multiple copies of the sequence within an individual) or quantitative number of molecules found to possess the polymorphism (e.g. determined by an INVADER assay) to be compared to the population allele frequency (or expected number), such that it is determined if the single individual is susceptible to a disease,

how far a disease has progressed (e.g. diseases such as cancer that may be diagnosed by identifying loss of heterozygosity), etc. In some embodiments, the individuals are from the same racial or ethnic class (e.g. European, African, Asian, Mexican, etc).

In particular embodiments, the present invention provides methods for detecting a rare mutation comprising; a) providing; i) a sample from a single subject, wherein the sample comprises at least 10,000 target nucleic acid sequences (e.g. from 10,000 cells, or at least 20,000 target nucleic acid sequences, or at least 100,000 target nucleic acid sequences), ii) a detection assay (e.g. the INVADER assay) capable of detecting a mutation in a population of target nucleic acid sequence that is present at an allele frequency of 1:1000 or less compared to wild type alleles; and b) assaying the sample with the detection assay under conditions such that the presence or absence of a rare mutation (e.g. one present at an allele frequency of 1:100, or 1:500, or 1:1000 or less compared to the wild type) is detected. In some embodiments, the target nucleic acid sequences are genomic (e.g. not polymerase chain reaction, or PCR, amplified, but directly from a cell). In other embodiments, the target nucleic acid sequences are amplified (*e.g.*, by PCR).

In some embodiments, the present invention provides methods for detecting a rare mutation comprising; a) providing: i) a sample from a single subject, wherein the sample comprises at least 10,000 target nucleic acid sequences, ii) a detection assay capable of detecting a mutation in a population of target nucleic acid sequence that is present at an allele frequency of 1:1000 or less compared to wild type alleles; and b) assaying the sample with the detection assay under conditions such that an allele frequency in the sample of a rare mutation is determined. In some embodiments, the subject's allele frequency is compared statistically to a known reference allele frequency (e.g. determined by the methods of the present invention or other methods), such that a diagnosis may be made (e.g. extent of disease, likelihood of having the disease, or passing it on to offspring, etc).

The present invention also provides methods for determining the number of molecules of one or more polymorphisms present in a sample by employing, for example, the INVADER assay (e.g. polymorphisms such as SNPs that are associated with disease). This assay may be used to determine the number of a particular polymorphism in a first sample, and then determining if there is a statistically significant difference between that number and the number

of the same polymorphism in a second sample. Preferably, one sample represents the number of the polymorphism expected to occur in a sample obtained from a healthy individual, or from a healthy population if pooled samples are used. A statistically significant difference between the number of a polymorphism expected to be at a single-base locus in a healthy individual and the number determined to be in a sample obtained from a patient is clinically indicative.

The present invention relates to detection assay panels comprising an array of different detection assays. The detection assays include assays for detecting mutations in nucleic acid molecules and for detecting gene expression levels. Assays find use, for example, in the identification of the genetic basis of phenotypes, including medically relevant phenotypes and in the development of diagnostic products, including clinical diagnostic products. The present invention also provides systems and methods for data storage, including data libraries and computer storage media comprising detection assay data.

For example, the present invention provides a panel comprising an array, wherein the array comprises a plurality of different assays (e.g., greater than about 50 different assays). In some preferred embodiments, the assays are substantially similar to at least one assay shown in figure 96. In some embodiments, the arrays comprise greater than about 100 different assays (e.g., 100, 101, 102, . . ., 130, . . ., 500, . . ., 1000, . . ., 10,000, . . ., 30,000, . . .). In some preferred embodiments, the assays comprise biplex assays. In other preferred embodiments, the assays comprise multiplex assays. In some embodiments, the array is a microarray. In some preferred embodiments, the assays are provided on a solid surface. For example, in some embodiments, the assays are provided on a microtiter plate.

In some preferred embodiments, the assays comprise nucleic acid detection assay. For example, in some embodiments, the assays detect polymorphisms (e.g., single-nucleotide polymorphisms in nucleic acids), including direct detection of genomic DNA (e.g., human genomic DNA).

The present invention also provides methods for using panels. For example, the present invention provides a method comprising: a) providing: i) a panel comprising an array, said array comprising a plurality of different assays (e.g., detection assays) and ii) a sample; and b) exposing the sample to the panel under conditions such that at least one of the assays detects the

presence of a target nucleic acid in the sample. Any of the panels or detection assays described herein may be used in the method..

The present invention also provides system and methods for developing clinical products based on information obtained from the use of the panels. Systems and methods are also provided for collecting, storing, and analyzing information obtained from use of the panels. For example, the present invention provides data libraries comprising data collected from detection assay testing. For example, in some embodiments, the data libraries contain data obtained from an assay similar to at least one assay shown in Figure 96. In some embodiments, the data libraries contain information obtained from greater than about 100 different assays (e.g., 100, 101, 102, . . ., 130, . . ., 500, . . .). In some embodiments, data libraries include test result data including, but not limited to, the presence or absence of a mutation in nucleic acid from a sample, allele frequency information, quantitation data, and disease correlation data. In some preferred embodiments, the data libraries also provide information correlated to the test result data including, but not limited to, an identity of a testing facility, detection assay components used to generate the data, other related detection assay components, reaction conditions, the identity of a user who requested the manufacture of the detection assay, date of detection assay use and/or testing, detection assay reliability information (e.g., determined the in silico methods of the present invention), information pertaining to the target sequence interrogated by the detection, information pertaining to clinical approval or requirements, and the like. In some embodiments, the present invention provides computer storage medium containing the above information and systems and methods for storing, accessing, and retrieving the information.

The present invention further provides methods for simultaneously detecting a plurality of polymorphisms (e.g., SNPs). For example, the present invention provides systems and methods for simultaneously detecting 100 or more polymorphism (100, . . ., 1000, . . ., 10,000, . . ., 100,000, . . .). In some embodiments, the plurality of polymorphisms are detected in a single reaction sample (e.g., in a multiplex reaction). In some embodiments, the polymorphisms are present in genomic DNA and target sequences containing a single polymorphism are amplified prior to detection of the polymorphisms. In some embodiments, the amplification comprises PCR amplification. In some embodiments, amplification is carried out such that there is a 10^5-10^6-fold increase in copies of the target sequence.

The present invention further provides system and methods for developing detection assays based on the design of a pre-validated detection assay. For example, the present invention provides thousands of specific INVADER detections assays directed at different target nucleic acid sequences, as well as components that find use in other detection assay formats. In some embodiments, one or more components of these assays are used in or are used in the design of a different type of detection assay. For example, validated target sequences may be used as targets in other types of detection assay. Likewise, oligonucleotides that hybridize to target sequences may be used directly, or in the design of hybridization oligonucleotides for other types of detection assays. The present invention is not limited in the nature of the detection assay that is produced using information from the thousands of INVADER detection assays (e.g., assays described in Figure 96). Such detection assays include, but are not limited to, hybridization methods and array technologies (e.g., Aclara BioSciences, Haywood, CA; Affymetrix, Santa Clara, CA; Agilent Technologies, Inc., Palo Alto, CA; Aviva Biosciences Corp., San Diego, CA; Caliper Technologies Corp., Palo Alto, CA; Celera, Rockville, MD; CuraGen Corp., New Haven, CT; Hyseq Inc., Sunnyvale, CA; Illumina, Inc., San Diego, CA; Incyte Genomics, Palo Alto, CA; Motorola BioChip Systems; Nanogen, San Diego, CA; Orchid BioSciences, Inc., Princeton, NJ; Applera Corp., Foster City, CA; Rosetta Inpharmatics, Kirkland, WA; and Sequenom, San Diego, CA); polymerase chain reaction; branched hybridization methods; enzyme mismatch cleavage methods; NASBA; sandwich hybridization methods; methods employing molecular beacons; ligase chain reactions, and the like.

The present invention relates to systems and methods for managing genetic information and medical records. For example, the present invention provides systems and methods for collecting, storing, and retrieving patient-specific genetic information from one or more electronic databases.

For example, in some embodiments, the present invention provides an electronic medical record comprising genetic information of a subject (e.g., single nucleotide polymorphism data of an animal or human patient) correlated to electronic medical history data of said subject. The present invention is not limited by the nature of the medical history data. Such data included, but is not limited to prescription data (e.g., data related to one or more drugs or other prescribed medical interventions of the subject, including drug identity, drug reaction data, allergies, risk

assessment data, and multi-drug interaction data, billing code levels, order restrictions); information pertaining a physician visit (e.g., date and time of visit, identity of physicians, physician notes, diagnosis information, differential diagnosis information, patient location, patient status, order status, referral information); patient identification information (e.g., patient age, gender, race, insurance carrier, allergies, past medical history, family history, social history, religion, employer, guarantor, address, contact information, patient condition code); and laboratory information (e.g., labs, radiology, and tests).

In some embodiments, the genetic information comprises single nucleotide polymorphism data (e.g., data related to the presence of one or more single nucleotide polymorphisms in the genetic material of the subject, including, but not limited to, the identity of the polymorphisms, the location of the polymorphisms, medical conditions associated with the presence or absence of the polymorphisms, detection assays information) and/or information related to single nucleotide polymorphism data (e.g., allele frequency of the polymorphism in one or more populations).

In some embodiments, the single nucleotide polymorphism data comprises data derived from an *in vitro* diagnostic single nucleotide polymorphism detection assay. In some embodiments, the single nucleotide polymorphism data comprises data derived from a panel comprising a plurality of single nucleotide polymorphism detection assays. In some preferred embodiments, the panel comprises a detection assays that detects medically associated single nucleotide polymorphisms (e.g., single nucleotide polymorphisms associated with a disease). In some embodiments, the detection assays detect polymorphisms associated with one or more medically relevant subject areas including, but not limited to cardiovascular disease, oncology, immunology, metabolic disorders, neurological disorders, musculoskeletal disorders, endocrinology, and genetic disease. In some embodiments, the panel comprises a plurality of single nucleotide polymorphism detection assays associated with two or more diseases. In some embodiments, the panel comprises a plurality of single nucleotide polymorphism detection assays that detect polymorphisms in drug metabolizing enzymes.

In some embodiments, the single nucleotide polymorphism data comprises data derived from a plurality of *in vitro* diagnostic single nucleotide polymorphism detection assays. In some embodiments, the detection assays comprises two or more unique invasive cleavage assays

(INVADER assay, Third Wave Technologies, Madison, WI). In some embodiments, one or more of the two or more unique invasive cleavage assays detected at least one single nucleotide polymorphism. In some embodiments, the single nucleotide polymorphism is associated with a medical condition. In some embodiments, the two or more unique invasive cleavage assays comprise at least 10 unique detection assays (e.g., 10, 11, 12, . . ., 100, . . ., 1000, . . ., 10,000, . . ., 50,000, . . .).

In some embodiments, the single nucleotide polymorphism data is derived from an analyte-specific reagent assay. In some embodiments, the single nucleotide polymorphism data is derived from at least one clinically valid detection assay.

The electronic medical records of the present invention may be located on any number of computers or devices. For example, in some embodiments, the electronic medical record is contained in a computer system of a patient, an insurance company, a health care provider (e.g., a physician, a hospital, a clinic, a health maintenance organization), a government agency, and a drug retailer or drug wholesaler, or pharmaceutical company. In some embodiments, the electronic medical record is stored on a small device to be carried on or in a subject (e.g., a personal digital assistant, a MED-ALERT bracelet, a smart card, and an implanted data storage device such as those described in U.S. Pat. No. 5,499,626, herein incorporated by reference in its entirety).

In some embodiments, the electronic medical record comprises addition information, including, but not limited to, medical billing data, insurance claim data, and scheduling data.

The present invention also provides a computer system comprising the electronic medical records described herein. In some embodiments, the computer system is configured for receiving data from the Internet (e.g., e.g., single nucleotide polymorphism data or one or more SNP assay(s) result data). In some embodiments, the computer system comprises one or more hardware or software components configured to carry out a processing routine. For example, in some embodiments, a software application is configured to receive single nucleotide polymorphism data automatically via a communications network. In other embodiments, the computer system comprises a routine for categorizing data (e.g., by disease type, by patient type, by genetic loci, etc.). In some embodiments, the computer system comprises a routine for carrying out a bioinformatics analysis routine (e.g., as described elsewhere herein). In some

embodiments, the computer system comprises a routine for carrying out a mathematical manipulation routine.

The present invention further provides a method for determining a correlation between a polymorphism (e.g., a SNP) and a phenotype, comprising: a) providing: samples from a plurality of subjects; medical records from the plurality of subjects, wherein the medical records contain information pertaining to a phenotype of the subjects; and detection assays that detect a polymorphism; b) exposing the samples to the detection assays under conditions such that the presence or absence of at least one polymorphism is revealed; and; c) determining a correlation between the at least one polymorphism and the phenotype of the subjects. In some embodiments, the plurality of subjects comprises 1000 or more subjects (e.g., 10,000 or more subjects). In some embodiments, the information pertaining to a phenotype comprises information pertaining to a disease. In other embodiments, the information pertaining to a phenotype comprises information pertaining to a drug interaction. In some embodiments, the medical record comprises an electronic medical record. While the present invention is not limited by the nature of the sample, in some preferred embodiments, the sample comprises a blood sample or a tissue biopsy.

The present invention also provides an electronic library comprising a plurality of electronic medical records for different subjects, each of the electronic medical records comprising, polymorphism data (e.g., single nucleotide polymorphism data) of the subject correlated to electronic medical history data of the subject. In some embodiments, the electronic medical history data comprises prescription data. In other embodiments, the prescription data comprises drug reaction data. In some embodiments, the single nucleotide polymorphism data comprises data derived from one or more *in vitro* diagnostic single nucleotide polymorphisms detection assays. In some embodiments, the single nucleotide polymorphism data comprises data derived from a panel, said panel comprising a plurality of single nucleotide polymorphisms detection assays. In some embodiments, the panel comprises detection assays that detect medically associated single nucleotide polymorphisms. In some embodiments, the panel comprises a plurality of single nucleotide polymorphisms detection assays that detect single nucleotide polymorphisms associated with a disease. In some embodiments, the panel comprises a plurality of detection assays that detect polymorphisms associated with one or more medically

relevant subject areas including, but not limited to, cardiovascular disease, oncology, immunology, metabolic disorders, neurological disorders, musculoskeletal disorders, endocrinology, and genetic disease. In some embodiments, the panel comprises a plurality of single nucleotide polymorphism detection assays associated with two or more diseases. In some embodiments, the panel comprises a plurality of single nucleotide polymorphism detection assays that detect polymorphisms in drug metabolizing enzymes. In some embodiments, the single nucleotide polymorphism data comprises data derived from a plurality of *in vitro* diagnostic single nucleotide polymorphism detection assays for each said different subject. In some embodiments, the detection assays comprises two or more unique invasive cleavage assays. In some embodiments, the one or more of the two or more unique invasive cleavage assays detected at least one single nucleotide polymorphism. In some preferred embodiments, the at least one single nucleotide polymorphism is associated with a medical condition.

The present invention is not limited by the number of unique invasive cleavage assays used in the method. In some embodiments, the two or more unique invasive cleavage assays comprise at least 10 unique detection assays (e.g., at least 1000, 10,000, 35,000, or more).

In some embodiments, the single nucleotide polymorphism data for each of the different subjects is derived from an analyte-specific reagent assay. In some embodiments, the single nucleotide polymorphism data for each of the different subjects is derived from at least one clinically valid detection assay.

The present invention also provides computer systems comprising the electronic libraries. In some embodiments, the computer system is configured for securely receiving single nucleotide polymorphism data from the Internet. In some embodiments, the computer system further comprises a routine to receive single nucleotide polymorphism data for each of the different subjects automatically via a communications network. In some embodiments, the computer system further comprises a routine to receive single nucleotide polymorphism data for each the different subjects from nodes of a national, regional or world-wide communications network. In some embodiments, the computer system further comprises a software application for categorizing the data for the different subjects. In some embodiments, the computer system further comprises a software application for carrying out a bioinformatics analysis on said data for each said different subject.

The present invention provides systems and methods for acquiring and analyzing biological information. In particular, the present invention provides systems and methods for developing detection assays and for use of detection assays in basic research discovery to facilitate selection and development of clinical detection assays.

In some embodiments, the present invention provides methods of validating a detection assay, comprising: a) collecting test result data from a plurality of users, wherein the test result data is generated with one or more detection panels, and wherein the detection panels comprise a plurality of candidate detection assays configured for target detection; and b) processing at least a portion of the test result data such that at least one valid detection assay is identified from the plurality of candidate detection assays. In other embodiments, the method further comprises step c) marketing said valid detection assay as an Analyte-Specific Reagent or an In-Vitro Diagnostic. In certain embodiments, said marketing comprises selling and/or advertising. In other embodiments, the present invention provides methods of validating a detection assay, comprising: a) distributing one or more detection panels to a plurality of users, wherein the detection panels comprise a plurality of candidate detection assays configured for target detection; b) collecting test result data from at least a portion of the plurality of users, wherein the test result data is generated with the detection panels; and c) processing at least a portion of the test result data such that at least one valid detection assay is identified from the plurality of candidate detection assays. In other embodiments, the method further comprises step d) marketing said valid detection assay as an Analyte-Specific Reagent or an In-Vitro Diagnostic. In certain embodiments, said marketing comprises selling and/or advertising.

In particular embodiments, the plurality of detection assays comprise two or more unique detection assays (e.g. 10, ... 50, ... 100, ... 1000, or more unique detection assays). In some embodiments, the plurality of detection assays comprise two or more unique INVADER assays (e.g. 10, ... 50, ... 100, ... 1000, or more unique INVADER assays).

In certain embodiments, the methods of the present invention further comprise a distribution system, wherein the distributing is accomplished with the distribution system. In some embodiments, the distributing one or more detection panels to the plurality of users is at a reduced cost. In other embodiments, the distributing one or more detection panels to the

plurality of users is at a subsidized cost. In still other embodiments, the distributing one or more detection panels to the plurality of users is at no cost.

In certain embodiments, prior to step a), the method further comprises the step of employing one or more of the plurality of candidate detection assays to discover at least one single nucleotide polymorphism. In particular embodiments, the plurality of detection assays comprise INVADER assays. In other embodiments, prior to step a), the method further comprises the step of utilizing one or more of the plurality of candidate detection assays to associate a single nucleotide polymorphism with a medical condition. In certain embodiments, the plurality of detection assays comprise INVADER assay components. In some embodiments, prior to step a), the method further comprises the step of utilizing one or more of the plurality of candidate detection assays, and computer aided analysis, to associate a single nucleotide polymorphism with a medical condition. In certain embodiments, the plurality of detection assays comprise INVADER assay components. In other embodiments, the INVADER assay components comprise an INVADER oligonucleotide, a probe, and a control target sequence. In particular embodiments, the plurality of detection assays comprise TAQMAN assay components (e.g. a probe and control target sequence).

In some embodiments, the one or more detection panels are configured for detecting a marker associated with a disease category. In certain embodiments, the disease category is selected from cardiovascular disease, cancer, autoimmune disease, metabolic disorders, neurological disease, musculoskeletal disorders, and endocrine related diseases.

In certain embodiments of the methods of the present invention, the plurality of users comprise researchers. In other embodiments, the plurality of users comprises at least 10 individual users. In some embodiments, the plurality of users comprises at least 200 individual users. In particular embodiments, the plurality of users comprises at least 500 individual users. In still other embodiments, the plurality of users comprises at least 1000 individual users. In particular embodiments, the plurality of users comprises at least 10,000 individual users.

In some embodiments of the methods of the present invention, the plurality of detection assays comprises at least 10 unique detection assays. In other embodiments, the plurality of detection assays comprises at least 1000 unique detection assays. In particular embodiments, the

plurality of detection assays comprises at least 10,000 unique detection assays. In certain embodiments, the plurality of detection assays comprises at least 50,000 unique detection assays.

In particular embodiments, the method further comprises a step, after the processing step, of selling the at least one valid detection assay as an Analyte Specific Reagent (ASR). In some embodiments the method further comprises a step, after the processing step, of selling the at least one valid detection assay as an Analyte Specific Reagent (ASR) to an In-Vitro Diagnostic Manufacturer or to a non-clinical laboratory. In additional embodiments, the method further comprises a step, after the processing step, of selling the at least one valid detection assay as an In-Vitro Diagnostic.

In some embodiments, the test result data comprises raw assay data. In other embodiments, test result data comprises analyzed assay data. In certain embodiments, the test result data comprises both raw assay data and analyzed assay data. In particular embodiments, the test result data comprises data resulting from testing of at least separate samples (e.g. at least 1000, at least 10,000, or at least 100,000 separate samples).

In certain embodiments, the collecting comprises receiving the test result data from at least a portion of the plurality of users over a communications network (e.g. Internet or World Wide Web). In some embodiments, the collecting further comprises storing the test result data in a database. In particular embodiments, the database is part of a computer system of a service provider. In certain embodiments, the collecting comprises receiving the test result data over the Internet. In some embodiments, the collecting comprises retrieving the test result data from a user's computer system over a communication network. In additional embodiments, the user's computer system comprises a software application configured to receive the test result data. In some embodiments, the software application is further configured to transmit the test result data automatically via a communications network.

In some embodiments, the processing comprises categorizing the test result data (e.g. arranging the data according to unique detection assay and/or type of medical condition associated with detection of a target). In other embodiments, the processing comprises in silico analysis. In certain embodiments, the processing comprises computer aided analysis of the test result data. In additional embodiments, the processing comprises mathematical manipulation of the test result data. In further embodiments, the processing comprises comparing the test result

data to a substantially equivalent predicate assay. In particular embodiments, the processing comprises mathematical manipulation of the test result data, and comparing the test result data to a substantially equivalent predicate assay.

In certain embodiments, at least one valid detection assay is identified as a result of being substantially equivalent to a predicate assay. In some embodiments, processing at least a portion of the test result data generates assay validation information.

In some embodiments, the methods of the present invention further comprise step e) submitting the assay validation information to a government body charged with approving products for clinical use. In certain embodiments, the government body is the Food and Drug Administration. In particular embodiments, the assay validation information is part of a 510(k) application that is submitted to the Food and Drug Administration. In other embodiments, the methods of the present invention further comprise a step of receiving approval from the Food and Drug Administration to market the at least one valid detection assay as an FDA approved In-Vitro diagnostic assay. In additional embodiments, the FDA approved In-Vitro diagnostic assay is a predicate for determining substantially equivalency for other In-Vitro diagnostic assays.

In some embodiments, the target is a single nucleotide polymorphism (e.g. in a DNA or RNA molecule). In other embodiments, the target is RNA (e.g. such that RNA expression can be quantitated).

The present invention also provides a method of developing an in-vitro diagnostic DNA or RNA analysis product comprising, running an assay through a product development funnel, in which the assay that enters the product development funnel is substantially similar to the in-vitro diagnostic DNA or RNA analysis product. In some embodiments, the assay is an assay to detect a single nucleotide polymorphism. In some preferred embodiments, the product development funnel optionally comprises one or more of the following: a discovery portion, a medically associated portion, an analyte-specific reagent portion, and an in-vitro diagnostic portion. In some embodiments, the assay comprises a chromosome specific assay. In some embodiments, the method further comprises the step of using a panel, wherein the panel comprises the assay. In other embodiments, the panel comprises a whole genome panel.

In some embodiments, the medically associated portion of the funnel comprises a panel organized by disease. In some preferred embodiments, the panel organized by disease is selected

from the group consisting of a cardiovascular disease panel, an oncology panel, an immunology panel, a metabolic disorders panel, a neurological disorders panel, a musculoskeletal disorders panel, an endocrinology panel, and a genetic disease panel.

In some embodiments, the method further comprises the step of using a panel, wherein the panel is a panel for a multiplicity of disease states and/or wherein the panel comprises a drug metabolizing enzyme panel.

The present invention further provides a method of increasing revenue and/or a profit margin from the development of an *in vitro* diagnostic DNA or RNA analysis product comprising channeling an assay through a product development funnel, in which the assay is substantially similar to the *in vitro* diagnostic DNA or RNA analysis product. In some embodiments, the *in vitro* DNA or RNA analysis product comprises an FDA approved product. In some preferred embodiments, the product development funnel has an ingress and an egress, wherein the assay is one of at least several thousand assays which enter the ingress. In other embodiments, the assay is one of about several hundred assays that exit the egress as the *in vitro* diagnostic DNA or RNA analysis product.

The present invention further provides a method of identifying single nucleotide polymorphisms comprising providing: 1) a plurality of samples comprising genomic DNA from a first individual and four or more additional individuals, each of the first and four or more additional individuals having genomic DNA comprising a first region, said first individual having a first single nucleotide polymorphism in the first region, 2) at least one detection reagent capable of generating a signal; and 3) at least one oligonucleotide probe designed to cause the detection reagent to generate a signal following contact of the probe with a portion of the first region of the genomic DNA of the first individual; contacting each of the genomic DNA samples with the oligonucleotide probe under conditions such that a signal is detected for the genomic DNA of the first individual; identifying at least one of the four or more additional individuals for which no signal is detected, thereby identifying a negative-tested individual; and assaying the first region of the negative-tested individual under conditions such that a second single nucleotide polymorphism is revealed in the first region of the genomic DNA of the negative-tested individual in addition to the first single nucleotide polymorphism, wherein the first individual lacks the second single nucleotide polymorphism. In some embodiments, the method

further provides a second oligonucleotide probe designed to cause the detection reagent to generate a signal following contact of the probe with a portion of the first region of the genomic DNA of the negative-tested individual, wherein the second oligonucleotide probes is contacted with the genomic DNA sample of the negative-tested individual. The second probe may be used concurrently with the first probe or may be used after the first probe (e.g., experiments conducted with the first probe may lead to the design of a second probe e.g., using the systems and methods of the present invention). The method may also include identifying negative detection assay results that are the result of one or more individuals lacking the first single nucleotide polymorphism.

DESCRIPTION OF THE FIGURES

The following figures form part of the present specification and are included to further demonstrate certain aspects and embodiments of the present invention. The invention may be better understood by reference to one or more of these figures in combination with the description of specific embodiments presented herein.

Figure 1 shows a general overview of the systems of the present invention.

Figured 2a-2f show various embodiments of INVADER LOCATOR computer interface displays.

Figure 3 shows an overview of in silico analysis in some embodiments of the present invention.

Figure 4 shows an overview of information flow for the design and production of detection assays in some embodiments of the present invention.

Figure 5 shows how the in silico processes of the present invention allow information to be processed to generate useful detection panels.

Figure 6 shows one embodiment of the INVADER detection assay.

Figure 7 shows a computer display of an INVADERCREATOR Order Entry screen.

Figure 8 shows a computer display of an INVADERCREATOR Multiple SNP Design Selection screen.

Figure 9 shows a computer display of an INVADERCREATOR Designer Worksheet screen.

Figure 10 shows a computer display of an INVADERCREATOR Output Page screen.

Figure 11 shows a computer display of an INVADERCREATOR Printer Ready Output screen.

Figure 12 A-12R show various SNP INVADER CREATOR (SIC) computer interface displays.

Figures 13A-13Q show various RIC INVADERCREATOR computer interface displays.

Figures 14a-14f show various TIC INVADER CREATOR computer interface displays.

Figure 15 shows an input target sequence and the result of processing this sequence with systems and routines of the present invention.

Figure 16 shows an example of a basic work flow for highly multiplexed PCR using the INVADER Medically Associated Panel.

Figure 17 shows a flow chart outlining the steps that may be performed in order to generate a primer set useful in multiplex PCR.

Figures 18-22 show sequences used and data generated in connection with PCR Primer Design Example 1.

Figures 23-30 show sequences used and data generated in connection with Example 2.

Figure 31 shows certain PCR primers useful for amplifying various regions of CYP2D6.

Figure 32 shows one protocol for Multiplex PCR optimization according to the present invention.

Figure 33 illustrates a perspective view of an exemplary synthesizer.

Figure 34 illustrates a cross-sectional view of an exemplary synthesizer.

Figure 35 illustrates a perspective view of a cartridge, chamber bowl and chamber seal of the present invention.

Figure 36 illustrates a detailed view of an exemplary cartridge.

Figure 37 illustrates an exemplary drain plate.

Figure 38A illustrates a top view of one embodiment of a drain plate. Figure 38B illustrates a top view of another embodiment of a drain plate gasket.

Figure 39 illustrates a side view of a drain plate gasket situated between a cartridge and a drain plate.

Figure 40 illustrates a cross-sectional view of a waste tube system.

Figure 41 illustrates a chamber bowl with chamber drain.

Figures 42A-C illustrate different embodiments of energy input components 95 and mixing components 96.

Figures 43A-B illustrate different combinations of energy input components 95 and mixing components 96.

Figure 44 illustrates one embodiment of a synthesis column.

Figure 45 illustrates a computer system coupled to a synthesizer.

Figures 46A-C illustrate 3 cross-sectional detailed views of different embodiments of a cartridge, drain plate, drain plate gasket, receiving hole of cartridge, and synthesis column.

Figure 47A and 47B illustrate embodiments of reagent dispense stations.

Figure 48A illustrates a synthesizer having a ventilation opening in a lid enclosure.

Figures 48B and 48C illustrate a synthesizer having ventilation tubing attached to a ventilation opening in a lid enclosure.

Figures 49A-C illustrate synthesizers having ventilated workspaces.

Figures 50A and 50B provide cross sectional views of an exemplary synthesizer having a lid enclosure 102, and illustrate air flow 109 toward the ventilation tubing 103 when the lid enclosure 102 is in a closed or opened position, respectively.

Figures 51A and 51B provide cross sectional views of an exemplary synthesizer having a primarily enclosed space in a base 2, and illustrate air flow 109 toward the ventilation tubing 103 when the lid enclosure 102 is in a closed or opened position, respectively.

Figure 52 illustrates a synthesizer 1, a robotic means 92, a cleave and deprotect component 93 and a purification component 94.

Figure 53 shows a schematic diagram of a polymer synthesizer of the present invention.

Figure 54A shows a side view of a reagent dispenser (2). Figure 54B shows a cross-sectional view of a reagent dispenser (2).

Figures 55A and 55B show a preferred embodiment of the reagent dispenser (2), wherein the outer surface of the delivery channel (9) contains first (13) and second (14) ring seals configured to form an airtight or substantially airtight seal with one or more points on the interior surface of a synthesis column (15) or other reaction chamber (e.g., with reaction chambers present in a synthesizer or a cleavage and deprotection component).

Figure 56 shows a solvent delivery component in one embodiment of the present invention.

Figure 57 shows a waste storage and purge component in one embodiment of the present invention.

Figured 58A-K show flow charts depicting the integrated data and process flows employed in the oligonucleotide production systems of the present invention.

Figure 59A-D show various protocols for high throughput, automated genotyping.

Figure 60A-60H various embodiments of the cleave and deprotect devices, and components thereof, of the present invention.

Figure 61 shows one embodiment of a data management system of the present invention.

Figure 62 shows another embodiment of a data management system of the present invention.

Figure 63 shows a computer display of an association database.

Figure 64 shows a computer display of a Microsoft Excel worksheet having data received by export from an association database.

Figure 65 shows a computer display of a plate viewer.

Figure 66 shows a computer display of a data viewer.

Figure 67 shows a computer display of allele caller results, having SNP results data displayed in the cells.

Figure 68 shows a computer display of allele caller results, having analyzed input assay data (in this example, a calculated ratio) displayed in the cells.

Figure 69 shows a computer display of a Microsoft Excel worksheet having SNP results data received by export from an allele caller.

Figure 70 shows a graph demonstrating the ability of the INVADER assay to detect mutations in the APOC4 gene in pooled samples.

Figure 71 shows a graph demonstrating the ability of the INVADER assay to detect mutations in the CFTR gene in pooled samples.

Figures 72-75 show graphs of the results of experiments described in Pooled Sample - Example 3.

Figure 76A shows data measuring allele signals in INVADER assays for detection of alleles comprising the indicated percentages of the number of copies of each locus.

Figure 76B shows an Excel graph comparing theoretical allele frequencies to allele frequencies calculated from the INVADER assay data shown in Figure 5A.

Figure 77 shows an Excel graph and data comparing actual and calculated allele frequencies for each of 8 SNP loci detected in pooled genomic DNA from 8 different individuals.

Figure 78 shows an Excel graph and data showing calculated allele frequencies compared to fold-over-zero minus 1 (FOZ-1) measurements for SNP locus 132505 in genomic DNAs having different mixtures of these alleles.

Figure 79 shows an Excel graph and data showing calculated allele frequencies compared to fold-over-zero minus 1 (FOZ-1) measurements for SNP locus 131534 in genomic DNAs having different mixtures of these alleles.

Figures 80A-80C show the sequences of the probes configured for use in the assays described in Pooled Sample - Example 4 and synthetic targets for each allele. "Y" indicates an amine blocking group. The polymorphism and the dye that will be detected for each probe, when used in the exemplary assay configurations described in Example 4, are indicated.

Figure 81 shows an overview of the integration of components of the systems and methods of the present invention.

Figure 82 shows identified p450 2D6 polymorphisms.

Figure 83 shows CYP2D6 specific PCR amplification.

Figure 84 depicts biplex signal detection using INVADER assays to detect CYP2D6.

Figures 85 and 86 show the results of an INVADER assay screen of 175 individuals for various CYP2D6 polymorphisms.

Figure 87 shows the minor allele frequency by population for various SNP consortium/Third Wave Technologies SNPs.

Figure 88 shows a schematic summary of the flow of detection assay development in the present invention from research products to clinical products.

Figure 89 shows a schematic summary of the discovery phase of the diagram shown in Figure 88.

Figure 90 shows a schematic summary of the development of potential clinical markers phase of the diagram shown in Figure 88.

Figure 91 shows exemplary detection assay products from each phase of the diagram shown in Figure 88.

Figure 92 shows business revenue generation from products from each phase of the diagram shown in Figure 88. The arrows showing revenue/margin per detection assay are not quantitative, but simply show a qualitative increase for each layer of the funnel.

Figure 93 shows a flow chart depicting a disease associated assay development process.

Figure 94 shows an overview of an ASR Fast Track Process.

Figure 95 shows a flow chart depicting a process for identifying "Super SNPs."

Figure 96 shows INVADER assay components for detecting polymorphisms in certain genes.

Figure 97A-97D shows various steps in the quality control assessment methods and protocols of the present invention.

Figure 98 shows a general overview of the oligonucleotide production and processing systems of the present invention.

DEFINITIONS

To facilitate an understanding of the present invention, a number of terms and phrases are defined below:

As used herein, the terms “solid support” or “support” refer to any material that provides a solid or semi-solid structure with which another material can be attached. Such materials include smooth supports (e.g., metal, glass, plastic, silicon, and ceramic surfaces) as well as textured and porous materials. Such materials also include, but are not limited to, gels, rubbers, polymers, and other non-rigid materials. Solid supports need not be flat. Supports include any type of shape including spherical shapes (e.g., beads). Materials attached to solid support may be attached to any portion of the solid support (e.g., may be attached to an interior portion of a porous solid support material). Preferred embodiments of the present invention have biological molecules such as nucleic acid molecules and proteins attached to solid supports. A biological material is “attached” to a solid support when it is associated with the solid support through a

non-random chemical or physical interaction. In some preferred embodiments, the attachment is through a covalent bond. However, attachments need not be covalent or permanent. In some embodiments, materials are attached to a solid support through a "spacer molecule" or "linker group." Such spacer molecules are molecules that have a first portion that attaches to the biological material and a second portion that attaches to the solid support. Thus, when attached to the solid support, the spacer molecule separates the solid support and the biological materials, but is attached to both.

As used herein, the term "derived from a different subject," such as samples or nucleic acids derived from a different subjects refers to a samples derived from multiple different individuals. For example, a blood sample comprising genomic DNA from a first person and a blood sample comprising genomic DNA from a second person are considered blood samples and genomic DNA samples that are derived from different subjects. A sample comprising five target nucleic acids derived from different subjects is a sample that includes at least five samples from five different individuals. However, the sample may further contain multiple samples from a given individual.

As used herein, the term "treating together," when used in reference to experiments or assays, refers to conducting experiments concurrently or sequentially, wherein the results of the experiments are produced, collected, or analyzed together (i.e., during the same time period). For example, a plurality of different target sequences located in separate wells of a multiwell plate or in different portions of a microarray are treated together in a detection assay where detection reactions are carried out on the samples simultaneously or sequentially and where the data collected from the assays is analyzed together.

The terms "assay data" and "test result data" as used herein refer to data collected from performance of an assay (*e.g.*, to detect or quantitate a gene, SNP or an RNA). Test result data may be in any form, *i.e.*, it may be raw assay data or analyzed assay data (*e.g.*, previously analyzed by a different process). Collected data that has not been further processed or analyzed is referred to herein as "raw" assay data (*e.g.*, a number corresponding to a measurement of signal, such as a fluorescence signal from a spot on a chip or a reaction vessel, or a number corresponding to measurement of a peak, such as peak height or area, as from, for example, a mass spectrometer, HPLC or capillary separation device), while assay data that has been

processed through a further step or analysis (*e.g.*, normalized, compared, or otherwise processed by a calculation) is referred to as "analyzed assay data" or "output assay data".

As used herein, the term "database" refers to collections of information (e.g., data) arranged for ease of retrieval, for example, stored in a computer memory. A "genomic information database" is a database comprising genomic information, including, but not limited to, polymorphism information (i.e., information pertaining to genetic polymorphisms), genome information (i.e., genomic information), linkage information (i.e., information pertaining to the physical location of a nucleic acid sequence with respect to another nucleic acid sequence, e.g., in a chromosome), and disease association information (i.e., information correlating the presence of or susceptibility to a disease to a physical trait of a subject, e.g., an allele of a subject). "Database information" refers to information to be sent to databases, stored in a database, processed in a database, or retrieved from a database. "Sequence database information" refers to database information pertaining to nucleic acid sequences. As used herein, the term "distinct sequence databases" refers to two or more databases that contain different information than one another. For example, the dbSNP and GenBank databases are distinct sequence databases because each contains information not found in the other.

As used herein, the terms "centralized control system" or "centralized control network" refer to information and equipment management systems (*e.g.*, a computer processor and computer memory) operable linked to a module or modules of equipment (*e.g.*, DNA synthesizers).

As used herein, the term "oligonucleotide synthesizer component" refers to a component of a system that is capable of synthesizing oligonucleotides (*e.g.*, a oligonucleotide synthesizers). In some embodiments, the oligonucleotide synthesizer component comprises a plurality of oligonucleotide synthesizers that are operably linked.

As used herein, the term "oligonucleotide processing component" refers to a component of a system capable of processing of oligonucleotides post-synthesis.. Examples of oligonucleotide processing stations include, but are not limited to, purification stations, dry-down stations, cleavage and deprotection stations, desalting stations, dilute and fill stations, and quality control stations.

As used herein, the terms "computer memory" and "computer memory device" refer to any storage media readable by a computer processor. Examples of computer memory include, but are not limited to, RAM, ROM, computer chips, digital video disc (DVDs), compact discs (CDs), hard disk drives (HDD), and magnetic tape.

As used herein, the term "computer readable medium" refers to any device or system for storing and providing information (*e.g.*, data and instructions) to a computer processor. Examples of computer readable media include, but are not limited to, DVDs, CDs, hard disk drives, magnetic tape and servers for streaming media over networks.

As used herein, the terms "processor" and "central processing unit" or "CPU" are used interchangeably and refers to a device that is able to read a program from a computer memory (*e.g.*, ROM or other computer memory) and perform a set of steps according to the program.

As used herein the term "oligonucleotide specification information" refers to any information used during the production of an oligonucleotide. Examples of oligonucleotide specification information includes, but is not limited to, sequence information, end-user (*e.g.*, customer) information, and concentration information (*e.g.*, the final concentration desired by the end-user).

As used herein the term "corresponding oligonucleotides" is used to refer to oligonucleotides that differ in at least one characteristic (*e.g.*, sequence, purity, required buffer, required salt concentration) and that are to be provided together (*e.g.*, in an INVADER assay, the INVADER oligonucleotide and Primary Probe are 'corresponding oligonucleotides').

As used herein, the term "divergent production" refers to the production of corresponding oligonucleotides employing at least two manufacturing stations, where a first corresponding oligonucleotide is never processed by at least one manufacturing station that is used to process a corresponding oligonucleotide.

As used herein the term "set of oligonucleotides" means at least two oligonucleotides that differ in at least one characteristic (*e.g.*, sequence, purity, required buffer, required salt concentration).

As used herein the term "purified sample," as in a purified oligonucleotide sample, refers to a sample where the full-length oligonucleotide in a sample is the predominate species of

oligonucleotide. For example, in some embodiments, at least 90%, preferably 95%, and more preferably 99% of oligonucleotides in a sample are full-length oligonucleotides.

As used herein, the terms "SNP," "SNPs" or "single nucleotide polymorphisms" refer to single base changes at a specific location in an organism's (*e.g.*, a human) genome. "SNPs" can be located in a portion of a genome that does not code for a gene. Alternatively, a "SNP" may be located in the coding region of a gene. In this case, the "SNP" may alter the structure and function of the RNA or the protein with which it is associated.

As used herein, the term "allele" refers to a variant form of a given sequence (e.g., including but not limited to, genes containing one or more SNPs). A large number of genes are present in multiple allelic forms in a population. A diploid organism carrying two different alleles of a gene is said to be heterozygous for that gene, whereas a homozygote carries two copies of the same allele.

As used herein, the term "linkage" refers to the proximity of two or more markers (e.g., genes) on a chromosome.

As used herein, the term "allele frequency" refers to the frequency of occurrence of a given allele (*e.g.*, a sequence containing a SNP) in given population (*e.g.*, a specific gender, race, or ethnic group). Certain populations may contain a given allele within a higher percent of its members than other populations. For example, a particular mutation in the breast cancer gene called BRCA1 was found to be present in one percent of the general Jewish population. In comparison, the percentage of people in the general U.S. population that have any mutation in BRCA1 has been estimated to be between 0.1 to 0.6 percent. Two additional mutations, one in the BRCA1 gene and one in another breast cancer gene called BRCA2, have a greater prevalence in the Ashkenazi Jewish population, bringing the overall risk for carrying one of these three mutations to 2.3 percent.

As used herein, the term "in silico analysis" refers to analysis performed using computer processors and computer memory. For example, "insilico SNP analysis" refers to the analysis of SNP data using computer processors and memory.

As used herein, the term "genotype" refers to the actual genetic make-up of an organism (e.g., in terms of the particular alleles carried at a genetic locus). Expression of the genotype gives rise to an organism's physical appearance and characteristics—the "phenotype."

As used herein, the term "locus" refers to the position of a gene or any other characterized sequence on a chromosome.

As used herein the term "disease" or "disease state" refers to a deviation from the condition regarded as normal or average for members of a species, and which is detrimental to an affected individual under conditions that are not inimical to the majority of individuals of that species (*e.g.*, diarrhea, nausea, fever, pain, and inflammation etc).

As used herein, the term "treatment" in reference to a medical course of action refer to steps or actions taken with respect to an affected individual as a consequence of a suspected, anticipated, or existing disease state, or wherein there is a risk or suspected risk of a disease state. Treatment may be provided in anticipation of or in response to a disease state or suspicion of a disease state, and may include, but is not limited to preventative, ameliorative, palliative or curative steps. The term "therapy" refers to a particular course of treatment.

The term "gene" refers to a nucleic acid (*e.g.*, DNA) sequence that comprises coding sequences necessary for the production of a polypeptide, RNA (e.g., rRNA, tRNA, etc.), or precursor. The polypeptide, RNA, or precursor can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired activity or functional properties (*e.g.*, ligand binding, signal transduction, etc.) of the full-length or fragment are retained. The term also encompasses the coding region of a structural gene and the including sequences located adjacent to the coding region on both the 5' and 3' ends for a distance of about 1 kb on either end such that the gene corresponds to the length of the full-length mRNA. The sequences that are located 5' of the coding region and which are present on the mRNA are referred to as 5' untranslated sequences. The sequences that are located 3' or downstream of the coding region and that are present on the mRNA are referred to as 3' untranslated sequences. The term "gene" encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region interrupted with non-coding sequences termed "introns" or "intervening regions" or "intervening sequences." Introns are segments included when a gene is transcribed into heterogeneous nuclear RNA (hnRNA); introns may contain regulatory elements such as enhancers. Introns are removed or "spliced out" from the nuclear or primary transcript; introns therefore are generally absent in the messenger RNA (mRNA) transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent

polypeptide. Variations (*e.g.*, mutations, SNPS, insertions, deletions) in transcribed portions of genes are reflected in, and can generally be detected in corresponding portions of the produced RNAs (*e.g.*, hnRNAs, mRNAs, rRNAs, tRNAs).

Where the phrase "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, amino acid sequence and like terms, such as polypeptide or protein are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

In addition to containing introns, genomic forms of a gene may also include sequences located on both the 5' and 3' end of the sequences that are present on the RNA transcript. These sequences are referred to as "flanking" sequences or regions (these flanking sequences are located 5' or 3' to the non-translated sequences present on the mRNA transcript). The 5' flanking region may contain regulatory sequences such as promoters and enhancers that control or influence the transcription of the gene. The 3' flanking region may contain sequences that direct the termination of transcription, post-transcriptional cleavage and polyadenylation.

The term "wild-type" refers to a gene or gene product that has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designed the "normal" or "wild-type" form of the gene. In contrast, the terms "modified," "mutant," and "variant" refer to a gene or gene product that displays modifications in sequence and or functional properties (*i.e.*, altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

As used herein, the terms "nucleic acid molecule encoding," "DNA sequence encoding," and "DNA encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these deoxyribonucleotides determines the order of amino acids along the polypeptide (protein) chain. In this case, the DNA sequence thus codes for the amino acid sequence.

DNA and RNA molecules are said to have "5' ends" and "3' ends" because mononucleotides are reacted to make oligonucleotides or polynucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor

in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotides or polynucleotide, referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide or polynucleotide, also may be said to have 5' and 3' ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5' of the "downstream" or 3' elements. This terminology reflects the fact that transcription proceeds in a 5' to 3' fashion along the DNA strand. The promoter and enhancer elements that direct transcription of a linked gene are generally located 5' or upstream of the coding region. However, enhancer elements can exert their effect even when located 3' of the promoter element and the coding region. Transcription termination and polyadenylation signals are located 3' or downstream of the coding region.

As used herein, the terms "an oligonucleotide having a nucleotide sequence encoding a gene" and "polynucleotide having a nucleotide sequence encoding a gene," means a nucleic acid sequence comprising the coding region of a gene or, in other words, the nucleic acid sequence that encodes a gene product. The coding region may be present in either a cDNA, genomic DNA, or RNA form. When present in a DNA form, the oligonucleotide or polynucleotide may be single-stranded (*i.e.*, the sense strand) or double-stranded. Suitable control elements such as enhancers/promoters, splice junctions, polyadenylation signals, *etc.* may be placed in close proximity to the coding region of the gene if needed to permit proper initiation of transcription and/or correct processing of the primary RNA transcript. Alternatively, the coding region utilized in the expression vectors of the present invention may contain endogenous enhancers/promoters, splice junctions, intervening sequences, polyadenylation signals, *etc.* or a combination of both endogenous and exogenous control elements.

As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (*i.e.*, a sequence of nucleotides) related by the base-pairing rules. For example, for the sequence "5'-A-G-T-3'," is complementary to the sequence "3'-T-C-A-5'." Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has

significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods that depend upon binding between nucleic acids.

The term "homology" refers to a degree of complementarity. There may be partial homology or complete homology (*i.e.*, identity). A partially complementary sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid and is referred to using the functional term "substantially homologous." The term "inhibition of binding," when used in reference to nucleic acid binding, refers to inhibition of binding caused by competition of homologous sequences for binding to a target sequence. The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (*i.e.*, the hybridization) of a completely homologous to a target under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (*i.e.*, selective) interaction. The absence of non-specific binding may be tested by the use of a second target that lacks even a partial degree of complementarity (*e.g.*, less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

The art knows well that numerous equivalent conditions may be employed to comprise low stringency conditions; factors such as the length and nature (DNA, RNA, base composition) of the probe and nature of the target (DNA, RNA, base composition, present in solution or immobilized, etc.) and the concentration of the salts and other components (*e.g.*, the presence or absence of formamide, dextran sulfate, polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of low stringency hybridization different from, but equivalent to, the above listed conditions. In addition, the art knows conditions that promote hybridization under conditions of high stringency (*e.g.*, increasing the temperature of the hybridization and/or wash steps, the use of formamide in the hybridization solution, etc.).

When used in reference to a double-stranded nucleic acid sequence such as a cDNA or genomic clone, the term "substantially homologous" refers to any probe that can hybridize to

either or both strands of the double-stranded nucleic acid sequence under conditions of low stringency as described above.

A gene may produce multiple RNA species that are generated by differential splicing of the primary RNA transcript. cDNAs that are splice variants of the same gene will contain regions of sequence identity or complete homology (representing the presence of the same exon or portion of the same exon on both cDNAs) and regions of complete non-identity (for example, representing the presence of exon "A" on cDNA 1 wherein cDNA 2 contains exon "B" instead). Because the two cDNAs contain regions of sequence identity they will both hybridize to a probe derived from the entire gene or portions of the gene containing sequences found on both cDNAs; the two splice variants are therefore substantially homologous to such a probe and to each other.

As used herein, the term "hybridization" is used in reference to the pairing of complementary nucleic acids. Hybridization and the strength of hybridization (*i.e.*, the strength of the association between the nucleic acids) is impacted by such factors as the degree of complementary between the nucleic acids, stringency of the conditions involved, the T_m of the formed hybrid, and the G:C ratio within the nucleic acids.

As used herein, the term "T_m" is used in reference to the "melting temperature." The melting temperature is the temperature at which a population of double-stranded nucleic acid molecules becomes half dissociated into single strands. The equation for calculating the T_m of nucleic acids is well known in the art. As indicated by standard references, a simple estimate of the T_m value may be calculated by the equation: $T_m = 81.5 + 0.41(\% G + C)$, when a nucleic acid is in aqueous solution at 1 M NaCl (*See e.g.*, Anderson and Young, Quantitative Filter Hybridization, *in Nucleic Acid Hybridization* [1985]). Other references include more sophisticated computations that take structural as well as sequence characteristics into account for the calculation of T_m.

As used herein the term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds such as organic solvents, under which nucleic acid hybridizations are conducted. Those skilled in the art will recognize that "stringency" conditions may be altered by varying the parameters just described either individually or in concert. With "high stringency" conditions, nucleic acid base pairing will occur only between nucleic acid fragments that have a high frequency of complementary base

sequences (*e.g.*, hybridization under "high stringency" conditions may occur between homologs with about 85-100% identity, preferably about 70-100% identity). With medium stringency conditions, nucleic acid base pairing will occur between nucleic acids with an intermediate frequency of complementary base sequences (*e.g.*, hybridization under "medium stringency" conditions may occur between homologs with about 50-70% identity). Thus, conditions of "weak" or "low" stringency are often required with nucleic acids that are derived from organisms that are genetically diverse, as the frequency of complementary sequences is usually less.

"High stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH_2PO_4 H_2O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5X Denhardt's reagent and 100 μg/ml denatured salmon sperm DNA followed by washing in a solution comprising 0.1X SSPE, 1.0% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

"Medium stringency conditions" when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH_2PO_4 H_2O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5X Denhardt's reagent and 100 μg/ml denatured salmon sperm DNA followed by washing in a solution comprising 1.0X SSPE, 1.0% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

"Low stringency conditions" comprise conditions equivalent to binding or hybridization at 42 C in a solution consisting of 5X SSPE (43.8 g/l NaCl, 6.9 g/l NaH_2PO_4 H_2O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.1% SDS, 5X Denhardt's reagent [50X Denhardt's contains per 500 ml: 5 g Ficoll (Type 400, Pharamcia), 5 g BSA (Fraction V; Sigma)] and 100 g/ml denatured salmon sperm DNA followed by washing in a solution comprising 5X SSPE, 0.1% SDS at 42 C when a probe of about 500 nucleotides in length is employed.

The following terms are used to describe the sequence relationships between two or more polynucleotides: "reference sequence," "sequence identity," "percentage of sequence identity," and "substantial identity." A "reference sequence" is a defined sequence used as a basis for a sequence comparison; a reference sequence may be a subset of a larger sequence, for example, as a segment of a full-length cDNA sequence given in a sequence listing or may comprise a

complete gene sequence. Generally, a reference sequence is at least 20 nucleotides in length, frequently at least 25 nucleotides in length, and often at least 50 nucleotides in length. Since two polynucleotides may each (1) comprise a sequence (*i.e.*, a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) may further comprise a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window," as used herein, refers to a conceptual segment of at least 20 contiguous nucleotide positions wherein a polynucleotide sequence may be compared to a reference sequence of at least 20 contiguous nucleotides and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by the local homology algorithm of Smith and Waterman [Smith and Waterman, *Adv. Appl. Math.* 2: 482 (1981)] by the homology alignment algorithm of Needleman and Wunsch [Needleman and Wunsch, *J. Mol. Biol.* 48:443 (1970)], by the search for similarity method of Pearson and Lipman [Pearson and Lipman, *Proc. Natl. Acad. Sci. (U.S.A.)* 85:2444 (1988)], by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by inspection, and the best alignment (*i.e.*, resulting in the highest percentage of homology over the comparison window) generated by the various methods is selected. The term "sequence identity" means that two polynucleotide sequences are identical (*i.e.*, on a nucleotide-by-nucleotide basis) over the window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

As applied to polynucleotides, the term "substantial identity" denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 85 percent sequence identity, preferably at least 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison window of at least 20 nucleotide positions, frequently over a window of at least 25-50 nucleotides, wherein the percentage of sequence identity is calculated by comparing the reference sequence to the polynucleotide sequence which may include deletions or additions which total 20 percent or less of the reference sequence over the window of comparison. The reference sequence may be a subset of a larger sequence, for example, as a splice variant of the full-length sequences.

As applied to polypeptides, the term "substantial identity" means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least 80 percent sequence identity, preferably at least 90 percent sequence identity, more preferably at least 95 percent sequence identity or more (*e.g.*, 99 percent sequence identity). Preferably, residue positions that are not identical differ by conservative amino acid substitutions. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

"Amplification" is a special case of nucleic acid replication involving template specificity. It is to be contrasted with non-specific template replication (*i.e.*, replication that is template-dependent but not dependent on a specific template). Template specificity is here distinguished from fidelity of replication (*i.e.*, synthesis of the proper polynucleotide sequence) and nucleotide (ribo- or deoxyribo-) specificity. Template specificity is frequently described in

terms of "target" specificity. Target sequences are "targets" in the sense that they are sought to be sorted out from other nucleic acid. Amplification techniques have been designed primarily for this sorting out.

Template specificity is achieved in most amplification techniques by the choice of enzyme. Amplification enzymes are enzymes that, under conditions they are used, will process only specific sequences of nucleic acid in a heterogeneous mixture of nucleic acid. For example, in the case of Q replicase, MDV-1 RNA is the specific template for the replicase (D.L. Kacian *et al.*, Proc. Natl. Acad. Sci. USA 69:3038 [1972]). Other nucleic acid will not be replicated by this amplification enzyme. Similarly, in the case of T7 RNA polymerase, this amplification enzyme has a stringent specificity for its own promoters (M. Chamberlin *et al.*, Nature 228:227 [1970]). In the case of T4 DNA ligase, the enzyme will not ligate the two oligonucleotides or polynucleotides, where there is a mismatch between the oligonucleotide or polynucleotide substrate and the template at the ligation junction (D.Y. Wu and R. B. Wallace, Genomics 4:560 [1989]). Finally, *Taq* and *Pfu* polymerases, by virtue of their ability to function at high temperature, are found to display high specificity for the sequences bounded and thus defined by the primers; the high temperature results in thermodynamic conditions that favor primer hybridization with the target sequences and not hybridization with non-target sequences (H.A. Erlich (ed.), *PCR Technology*, Stockton Press [1989]).

As used herein, the term "amplifiable nucleic acid" is used in reference to nucleic acids that may be amplified by any amplification method. It is contemplated that "amplifiable nucleic acid" will usually comprise "sample template."

As used herein, the term "sample template" refers to nucleic acid originating from a sample that is analyzed for the presence of "target" (defined below). In contrast, "background template" is used in reference to nucleic acid other than sample template that may or may not be present in a sample. Background template is most often inadvertent. It may be the result of carryover, or it may be due to the presence of nucleic acid contaminants sought to be purified away from the sample. For example, nucleic acids from organisms other than those to be detected may be present as background in a test sample.

As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting

as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (*i.e.*, in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). The primer is preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, source of primer and the use of the method.

As used herein, the term "probe" or "hybridization probe" refers to an oligonucleotide (*i.e.*, a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced synthetically, recombinantly or by PCR amplification, that is capable of hybridizing, at least in part, to another oligonucleotide of interest. A probe may be single-stranded or double-stranded. Probes are useful in the detection, identification and isolation of particular sequences. In some preferred embodiments, probes used in the present invention will be labeled with a "reporter molecule," so that is detectable in any detection system, including, but not limited to enzyme (*e.g.*, ELISA, as well as enzyme-based histochemical assays), fluorescent, radioactive, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label.

As used herein, the term "target" refers to a nucleic acid sequence or structure to be detected or characterized.

As used herein, the term "polymerase chain reaction" ("PCR") refers to the method of K.B. Mullis (*See e.g.*, U.S. Patent Nos. 4,683,195, 4,683,202, and 4,965,188, hereby incorporated by reference), which describe a method for increasing the concentration of a segment of a target sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the target sequence consists of introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective strands of the double stranded target sequence. To effect

amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the target molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing, and polymerase extension can be repeated many times (*i.e.*, denaturation, annealing and extension constitute one "cycle"; there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired target sequence. The length of the amplified segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of the repeating aspect of the process, the method is referred to as the "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the target sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified."

With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (*e.g.*, hybridization with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; incorporation of ^{32}P-labeled deoxynucleotide triphosphates, such as dCTP or dATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide or polynucleotide sequence can be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications.

As used herein, the terms "PCR product," "PCR fragment," and "amplification product" refer to the resultant mixture of compounds after two or more cycles of the PCR steps of denaturation, annealing and extension are complete. These terms encompass the case where there has been amplification of one or more segments of one or more target sequences.

As used herein, the term "amplification reagents" refers to those reagents (deoxyribonucleotide triphosphates, buffer, etc.), needed for amplification except for primers, nucleic acid template, and the amplification enzyme. Typically, amplification reagents along with other reaction components are placed and contained in a reaction vessel (test tube, microwell, etc.).

As used herein, the term "recombinant DNA molecule" as used herein refers to a DNA molecule that is comprised of segments of DNA joined together by means of molecular biological techniques.

As used herein, the term "antisense" is used in reference to RNA sequences that are complementary to a specific RNA sequence (*e.g.*, mRNA). The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. The designation (-) (*i.e.*, "negative") is sometimes used in reference to the antisense strand, with the designation (+) sometimes used in reference to the sense (*i.e.*, "positive") strand.

The term "isolated" when used in relation to a nucleic acid, as in "an isolated oligonucleotide" or "isolated polynucleotide" refers to a nucleic acid sequence that is identified and separated from at least one contaminant nucleic acid with which it is ordinarily associated in its natural source. Isolated nucleic acid is present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids are nucleic acids such as DNA and RNA found in the state they exist in nature. For example, a given DNA sequence (*e.g.*, a gene) is found on the host cell chromosome in proximity to neighboring genes; RNA sequences, such as a specific mRNA sequence encoding a specific protein, are found in the cell as a mixture with numerous other mRNAs that encode a multitude of proteins. However, isolated nucleic acids encoding a polypeptide include, by way of example, such nucleic acid in cells ordinarily expressing the polypeptide where the nucleic acid is in a chromosomal location different from that of natural cells, or is otherwise flanked by a different nucleic acid sequence than that found in nature. The isolated nucleic acid, oligonucleotide, or polynucleotide may be present in single-stranded or double-stranded form. When an isolated nucleic acid, oligonucleotide or polynucleotide is to be utilized to express a protein, the oligonucleotide or polynucleotide will contain at a minimum the sense or coding strand (*i.e.*, the oligonucleotide or polynucleotide may single-stranded), but may contain both the sense and anti-sense strands (*i.e.*, the oligonucleotide or polynucleotide may be double-stranded).

As used herein the term "portion" when in reference to a nucleotide sequence (as in "a portion of a given nucleotide sequence") refers to fragments of that sequence. The fragments may range in size from four nucleotides to the entire nucleotide sequence minus one nucleotide (e.g., 10 nucleotides, 11, . . ., 20, . . .).

As used herein, the term "purified" or "to purify" refers to the removal of contaminants from a sample. As used herein, the term "purified" refers to molecules (e.g., nucleic or amino acid sequences) that are removed from their natural environment, isolated or separated. An "isolated nucleic acid sequence" is therefore a purified nucleic acid sequence. "Substantially purified" molecules are at least 60% free, preferably at least 75% free, and more preferably at least 90% free from other components with which they are naturally associated.

The term "recombinant protein" or "recombinant polypeptide" as used herein refers to a protein molecule that is expressed from a recombinant DNA molecule.

The term "native protein" as used herein to indicate that a protein does not contain amino acid residues encoded by vector sequences; that is the native protein contains only those amino acids found in the protein as it occurs in nature. A native protein may be produced by recombinant means or may be isolated from a naturally occurring source.

As used herein the term "portion" when in reference to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four consecutive amino acid residues to the entire amino acid sequence minus one amino acid.

The term "Southern blot," refers to the analysis of DNA on agarose or acrylamide gels to fractionate the DNA according to size followed by transfer of the DNA from the gel to a solid support, such as nitrocellulose or a nylon membrane. The immobilized DNA is then probed with a labeled probe to detect DNA species complementary to the probe used. The DNA may be cleaved with restriction enzymes prior to electrophoresis. Following electrophoresis, the DNA may be partially depurinated and denatured prior to or during transfer to the solid support. Southern blots are a standard tool of molecular biologists (J. Sambrook *et al., Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, NY, pp 9.31-9.58 [1989]).

The term "Western blot" refers to the analysis of protein(s) (or polypeptides) immobilized onto a support such as nitrocellulose or a membrane. The proteins are run on acrylamide gels to separate the proteins, followed by transfer of the protein from the gel to a solid support, such as nitrocellulose or a nylon membrane. The immobilized proteins are then exposed to antibodies with reactivity against an antigen of interest. The binding of the antibodies may be detected by various methods, including the use of labeled antibodies.

The term "test compound" refers to any chemical entity, pharmaceutical, drug, and the like that are tested in an assay (*e.g.*, a drug screening assay) for any desired activity (*e.g.*, including but not limited to, the ability to treat or prevent a disease, illness, sickness, or disorder of bodily function, or otherwise alter the physiological or cellular status of a sample). Test compounds comprise both known and potential therapeutic compounds. A test compound can be determined to be therapeutic by screening using the screening methods of the present invention. A "known therapeutic compound" refers to a therapeutic compound that has been shown (*e.g.*, through animal trials or prior experience with administration to humans) to be effective in such treatment or prevention.

The term "sample" as used herein is used in its broadest sense. A sample suspected of containing a human chromosome or sequences associated with a human chromosome may comprise a cell, chromosomes isolated from a cell (*e.g.*, a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern blot analysis), RNA (in solution or bound to a solid support such as for Northern blot analysis), cDNA (in solution or bound to a solid support) and the like. A sample suspected of containing a protein may comprise a cell, a portion of a tissue, an extract containing one or more proteins and the like.

The term "label" as used herein refers to any atom or molecule that can be used to provide a detectable (preferably quantifiable) effect, and that can be attached to a nucleic acid or protein. Labels include but are not limited to dyes; radiolabels such as ^{32}P; binding moieties such as biotin; haptens such as digoxgenin; luminogenic, phosphorescent or fluorogenic moieties; and fluorescent dyes alone or in combination with moieties that can suppress or shift emission spectra by fluorescence resonance energy transfer (FRET). Labels may provide signals detectable by fluorescence, radioactivity, colorimetry, gravimetry, X-ray diffraction or absorption, magnetism, enzymatic activity, and the like. A label may be a charged moiety (positive or negative charge) or alternatively, may be charge neutral. Labels can include or consist of nucleic acid or protein sequence, so long as the sequence comprising the label is detectable.

The term "signal" as used herein refers to any detectable effect, such as would be caused or provided by a label or an assay reaction.

As used herein, the term "detector" refers to a system or component of a system, *e.g.*, an instrument (*e.g.* a camera, fluorimeter, charge-coupled device, scintillation counter, etc) or a reactive medium (X-ray or camera film, pH indicator, etc.), that can convey to a user or to another component of a system (*e.g.*, a computer or controller) the presence of a signal or effect. A detector can be a photometric or spectrophotometric system, which can detect ultraviolet, visible or infrared light, including fluorescence or chemiluminescence; a radiation detection system; a spectroscopic system such as nuclear magnetic resonance spectroscopy, mass spectrometry or surface enhanced Raman spectrometry; a system such as gel or capillary electrophoresis or gel exclusion chromatography; or other detection system known in the art, or combinations thereof.

As used herein, the term "distribution system" refers to systems capable of transferring and/or delivering materials from one entity to another or one location to another. For example, a distribution system for transferring detection panels from a manufacturer or distributor to a user may comprise, but is not limited to, a packaging department, a mail room, and a mail delivery system. Alternately, the distribution system may comprise, but is not limited to, one or more delivery vehicles and associated delivery personnel, a display stand, and a distribution center. In some embodiments of the present invention interested parties (*e.g.*, detection panel manufactures) utilize a distribution system to transfer detection panels to users at no cost, at a subsidized cost, or at a reduced cost.

As used herein, the term "at a reduced cost" refers to the transfer of goods or services at a reduced direct cost to the recipient (e.g. user). In some embodiments, "at a reduced cost" refers to transfer of goods or services at no cost to the recipient.

As used herein, the term "at a subsidized cost" refers to the transfer of goods or services, wherein at least a portion of the recipient's cost is deferred or paid by another party. In some embodiments, "at a subsidized cost" refers to transfer of goods or services at no cost to the recipient.

As used herein, the term "at no cost" refers to the transfer of goods or services with no direct financial expense to the recipient. For example, when detection panels are provided by a manufacturer or distributor to a user (e.g. research scientist) at no cost, the user does not directly pay for the tests.

The term "detection" as used herein refers to quantitatively or qualitatively identifying an analyte (*e.g.*, DNA, RNA or a protein) within a sample. The term "detection assay" as used herein refers to a kit, test, or procedure performed for the purpose of detecting an analyte nucleic acid within a sample. Detection assays produce a detectable signal or effect when performed in the presence of the target analyte, and include but are not limited to assays incorporating the processes of hybridization, nucleic acid cleavage (*e.g.*, exo- or endonuclease), nucleic acid amplification, nucleotide sequencing, primer extension, or nucleic acid ligation.

As used herein, the term "functional detection oligonucleotide" refers to an oligonucleotide that is used as a component of a detection assay, wherein the detection assay is capable of successfully detecting (i.e., producing a detectable signal) an intended target nucleic acid when the functional detection oligonucleotide provides the oligonucleotide component of the detection assay. This is in contrast to a non-functional detection oligonucleotides, which fail to produce a detectable signal in a detection assay for the particular target nucleic acid when the non-functional detection oligonucleotide is provided as the oligonucleotide component of the detection assay. Determining if an oligonucleotide is a functional oligonucleotide can be carried out experimentally by testing the oligonucleotide in the presence of the particular target nucleic acid using the detection assay.

As used herein, the term "hyperlink" refers to a navigational link from one document to another, or from one portion (or component) of a document to another. Typically, a hyperlink is displayed as a highlighted word or phrase that can be selected by clicking on it using a mouse to jump to the associated document or documented portion.

As used herein, the term "hypertext system" refers to a computer-based informational system in which documents (and possibly other types of data entities) are linked together via hyperlinks to form a user-navigable "web."

As used herein, the term "Internet" refers to any collection of networks using standard protocols. For example, the term includes a collection of interconnected (public and/or private) networks that are linked together by a set of standard protocols (such as TCP/IP, HTTP, and FTP) to form a global, distributed network. While this term is intended to refer to what is now commonly known as the Internet, it is also intended to encompass variations that may be made in the future, including changes and additions to existing standard protocols or integration with

other media (*e.g.*, television, radio, etc). The term is also intended to encompass non-public networks such as private (e.g., corporate) Intranets.

As used herein, the terms "World Wide Web" or "web" refer generally to both (i) a distributed collection of interlinked, user-viewable hypertext documents (commonly referred to as Web documents or Web pages) that are accessible via the Internet, and (ii) the client and server software components which provide user access to such documents using standardized Internet protocols. Currently, the primary standard protocol for allowing applications to locate and acquire Web documents is HTTP, and the Web pages are encoded using HTML. However, the terms "Web" and "World Wide Web" are intended to encompass future markup languages and transport protocols that may be used in place of (or in addition to) HTML and HTTP.

As used herein, the term "web site" refers to a computer system that serves informational content over a network using the standard protocols of the World Wide Web. Typically, a Web site corresponds to a particular Internet domain name and includes the content associated with a particular organization. As used herein, the term is generally intended to encompass both (i) the hardware/software server components that serve the informational content over the network, and (ii) the "back end" hardware/software components, including any non-standard or specialized components, that interact with the server components to perform services for Web site users.

As used herein, the term "HTML" refers to HyperText Markup Language that is a standard coding convention and set of codes for attaching presentation and linking attributes to informational content within documents. HTML is based on SGML, the Standard Generalized Markup Language. During a document authoring stage, the HTML codes (referred to as "tags") are embedded within the informational content of the document. When the Web document (or HTML document) is subsequently transferred from a Web server to a browser, the codes are interpreted by the browser and used to parse and display the document. Additionally, in specifying how the Web browser is to display the document, HTML tags can be used to create links to other Web documents (commonly referred to as "hyperlinks").

As used herein, the term "XML" refers to Extensible Markup Language, an application profile that, like HTML, is based on SGML. XML differs from HTML in that: information providers can define new tag and attribute names at will; document structures can be nested to any level of complexity; any XML document can contain an optional description of its grammar

for use by applications that need to perform structural validation. XML documents are made up of storage units called entities, which contain either parsed or unparsed data. Parsed data is made up of characters, some of which form character data, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. XML provides a mechanism to impose constraints on the storage layout and logical structure, to define constraints on the logical structure and to support the use of predefined storage units. A software module called an XML processor is used to read XML documents and provide access to their content and structure.

As used herein, the term "HTTP" refers to HyperText Transport Protocol that is the standard World Wide Web client-server protocol used for the exchange of information (such as HTML documents, and client requests for such documents) between a browser and a Web server. HTTP includes a number of different types of messages that can be sent from the client to the server to request different types of server actions. For example, a "GET" message, which has the format GET, causes the server to return the document or file located at the specified URL.

As used herein, the term "URL" refers to Uniform Resource Locator that is a unique address that fully specifies the location of a file or other resource on the Internet. The general format of a URL is protocol://machine address:port/path/filename. The port specification is optional, and if none is entered by the user, the browser defaults to the standard port for whatever service is specified as the protocol. For example, if HTTP is specified as the protocol, the browser will use the HTTP default port of 80.

As used herein, the term "PUSH technology" refers to an information dissemination technology used to send data to users over a network. In contrast to the World Wide Web (a "pull" technology), in which the client browser must request a Web page before it is sent, PUSH protocols send the informational content to the user computer automatically, typically based on information pre-specified by the user.

As used herein, the term "communication network" refers to any network that allows information to be transmitted from one location to another. For example, a communication network for the transfer of information from one computer to another includes any public or private network that transfers information using electrical, optical, satellite transmission, and the like. Two or more devices that are part of a communication network such that they can directly

or indirectly transmit information from one to the other are considered to be “in electronic communication” with one another. A computer network containing multiple computers may have a central computer (“central node”) that processes information to one or more sub-computers that carry out specific tasks (“sub-nodes”). Some networks comprises computers that are in “different geographic locations” from one another, meaning that the computers are located in different physical locations (i.e., aren’t physically the same computer, e.g., are located in different countries, states, cities, rooms, etc.).

As used herein, the term "detection assay component" refers to a component of a system capable of performing a detection assay. Detection assay components include, but are not limited to, hybridization probes, buffers, and the like.

As used herein, the term “a detection assay configured for target detection” refers to a collection of assay components that are capable of producing a detectable signal when carried out using the target nucleic acid. For example, a detection assay that has empirically been demonstrated to detect a particular single nucleotide polymorphism is considered a detection assay configured for target detection.

As used herein, the phrase "unique detection assay" refers to a detection assay that has a different collection of detection assay components in relation to other detection assays located on the same detection panel. A unique assay doesn't necessarily detect a different target (e.g. SNP) than other assays on the same detection panel, but it does have a least one difference in the collection of components used to detect a given target (e.g. a unique detection assay may employ a probe sequences that is shorter or longer in length than other assays on the same detection panel).

As used herein, the term "candidate" refers to an assay or analyte, *e.g.*, a nucleic acid, suspected of having a particular feature or property. A "candidate sequence" refers to a nucleic acid suspected of comprising a particular sequence, while a "candidate oligonucleotide" refers to an oligonucleotide suspected of having a property such as comprising a particular sequence, or having the capability to hybridize to a target nucleic acid or to perform in a detection assay. A "candidate detection assay" refers to a detection assay that is suspected of being a valid detection assay.

As used herein, the term "detection panel" refers to a substrate or device containing at least two unique candidate detection assays configured for target detection.

As used herein, the term "valid detection assay" refers to a detection assay that has been shown to accurately predict an association between the detection of a target and a phenotype (e.g. medical condition). Examples of valid detection assays include, but are not limited to, detection assays that, when a target is detected, accurately predict the phenotype medical 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, or 99.9% of the time. Other examples of valid detection assays include, but are not limited to, detection assays that quality as and/or are marketed as Analyte-Specific Reagents (*i.e.* as defined by FDA regulations) or In-Vitro Diagnostics (*i.e.* approved by the FDA).

As used herein, the term "kit" refers to any delivery system for delivering materials. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., oligonucleotides, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. As used herein, the term "fragmented kit" refers to a delivery systems comprising two or more separate containers that each contain a subportion of the total kit components. The containers may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second container contains oligonucleotides. The term "fragmented kit" is intended to encompass kits containing Analyte specific reagents (ASR's) regulated under section 520(e) of the Federal Food, Drug, and Cosmetic Act, but are not limited thereto. Indeed, any delivery system comprising two or more separate containers that each contains a subportion of the total kit components are included in the term "fragmented kit." In contrast, a "combined kit" refers to a delivery system containing all of the components of a reaction assay in a single container (e.g., in a single box housing each of the desired components). The term "kit" includes both fragmented and combined kits.

As used herein, the term "information" refers to any collection of facts or data. In reference to information stored or processed using a computer system(s), including but not limited to internets, the term refers to any data stored in any format (e.g., analog, digital, optical,

etc.). As used herein, the term "information related to a subject" refers to facts or data pertaining to a subject (e.g., a human, plant, or animal). The term "genomic information" refers to information pertaining to a genome including, but not limited to, nucleic acid sequences, genes, allele frequencies, RNA expression levels, protein expression, phenotypes correlating to genotypes, etc. "Allele frequency information" refers to facts or data pertaining allele frequencies, including, but not limited to, allele identities, statistical correlations between the presence of an allele and a characteristic of a subject (e.g., a human subject), the presence or absence of an allele in a individual or population, the percentage likelihood of an allele being present in an individual having one or more particular characteristics, etc.

As used herein, the term "assay validation information" refers to genomic information and/or allele frequency information resulting from processing of test result data (*e.g.* processing with the aid of a computer). Assay validation information may be used, for example, to identify a particular candidate detection assay as a valid detection assay.

As used herein, the term "coupled," as in "coupled attachment," refers to attachments between objects that do not, by themselves, provide a pressure-tight seal. For example, two metal plates that are attached by screws or pins may comprise a coupled attachment. While the two plates are attached, the seam between them does not form a pressure-tight seal (*i.e.*, gas and/or liquid can escape through the seam).

As used herein, the term "synthesis and purge component" refers to a component of a synthesizer containing a cartridge for holding one or more synthesis columns attached to or connected to a drain plate for allowing waste or wash material from the synthesis columns to be directed to a waste disposal system.

As used herein, the term "cartridge" refers to a device for holding one or more synthesis columns. For example, cartridges can contain a plurality of openings (*e.g.*, receiving holes) into which synthesis columns may be placed. "Rotary cartridges" refer to cartridges that, in operation, can rotate with respect to an axis, such that a synthesis column is moved from one location in a plane (a reagent dispensing location) to another location in the plane (a non-reagent dispensing location) following rotation of the cartridge.

As used herein, the term "nucleic acid synthesis column" or "synthesis column" refers to a container or chamber in which nucleic acid synthesis reactions are carried out. For example,

synthesis columns include plastic cylindrical columns and pipette tip formats, containing openings at the top and bottom ends. The containers may contain or provide one or more matrices, solid supports, and/or synthesis reagents necessary to carry out chemical synthesis of nucleic acids. For example, in some embodiments of the present invention, synthesis columns contain a solid support matrix on which a growing nucleic acid molecule may be synthesized. Nucleic acid synthesis columns may be provided individually; alternatively, several synthesis columns may be provided together as a unit, e.g., in a strip or array, or as device such as a plate having a plurality of suitable chambers. Columns may be constructed of any material or combination of materials that do not adversely affect (e.g., chemically) the synthesis reaction or the use of the synthesized product. For example, columns or chambers may comprise polymers such as polypropylene, fluoropolymers such as TEFLON, metals and other materials that are substantially inert to synthesis reaction conditions, such as stainless steel, gold, silicon and glass. In some embodiments, chambers comprise a coating of such a suitable material over a structure comprising a different material.

As used herein, the term "seal" refers to any means for preventing the flow of gas or liquid through an opening. For example, a seal may be formed between two contacted materials using grease, o-rings, gaskets, and the like. In some embodiments, one or both of the contacted materials comprises an integral seal, such as, *e.g.*, a ridge, a lip or another feature configured to provide a seal between said contacted materials. An "airtight seal" or "pressure tight seal" is a seal that prevents detectable amounts of air from passing through an opening. A "substantially airtight" seal is a seal that prevents all but negligible amounts of air from passing through an opening. Negligible amounts of air are amounts that are tolerated by the particular system, such that desired system function is not compromised. For example, a seal in a nucleic acid synthesizer is considered substantially airtight if it prevents gas leaks in a reaction chamber, such that the gas pressure in the reaction chamber is sufficient to purge liquid in synthesis columns contained in the reaction chamber following a synthesis reaction. If gas pressure is depleted by a leak such that synthesis columns are not purged (*e.g.*, resulting in overflow during subsequent synthesis rounds), then the seal is not a substantially airtight seal. A substantially airtight seal can be detected empirically by carrying out synthesis and checking for failures (e.g., column overflows) during one or a series of reactions.

As used herein, the term "sealed contact point" refers to sealed seams between two or more objects. Seals on sealed contact points can be of any type that prevent the flow of gas or liquid through an opening. For example, seals can sit on the surface of a seam (*e.g.*, a face seal) or can be placed within a seam, such that a circumferential contact is created within the seam.

As used herein, the term "alignment detector" refers to any means for detecting the position of an object with respect to another object or with respect to the detector. For example, alignment detectors may detect the alignment of a dispensing end of a dispensing device (*e.g.*, a reagent tube, a waste tube, etc.) to a receiving device (e.g., a synthesis column, a waste valve, etc.). Alignment detectors may also detect the tilt angle of an object (*e.g.*, the angle of a plane of an object with respect to a reference plane). For example, the tilt angle of a plate mounted on a shaft may be detected to ensure a proper perpendicular relationship between the plate and the shaft. Alignment detectors include, but are not limited to, motion sensors, infra-red or LED-based detectors, and the like.

As used herein, the term "alignment markers" refers to reference points on an object that allow the object to be aligned to one or more other objects. Alignment markers include pictorial markings (*e.g.*, arrows, dots, etc.) and reflective markings, as well as pins, raised surfaces, holes, magnets, and the like.

As used herein, the term "motor connector" refers to any type of connection between a motor and another object. For example a motor designed to rotate another object may be connected to the object through a metal shaft, such that the rotation of the shaft, rotates the object. The metal shaft would be considered a motor connector.

As used herein, the term "packing material" refers to material placed in a passageway (*e.g.*, a synthesis column) in a manner such that it provides resistance against a pressure differential between the two ends of the passageway (*i.e.* hinders the discharge of the pressure differential). Packing material may comprise a single material or multiple materials. For example, in some embodiments of the present invention, packing material comprising a nucleic acid synthesis matrix (*e.g.*, a solid support for nucleic acid synthesis such as controlled pore glass, polystyrene, etc.) and/or one or more frits are used in synthesis columns to maintain a pressure differential between the two ends of the synthesis column. Packing material may be distributed into the reaction chambers in a variety of forms. For example, synthesis support

matrix may be provided as a granular powder. In some embodiments, support matrix may be provided in a "pill" form, wherein an appropriate amount of a support material is held together with a binder to form a pill, and wherein one or more pills are provided to a reaction chamber, as appropriate for the scale of the intended reaction, and further wherein the binder is removed or inactivated (*e.g.*, during a wash step) to allow the powdered matrix to function in the same manner as an unbound powder. The use of a pill embodiment provides the advantages of facilitating the process of pre-measuring synthesis support materials, allowing easy storage of support matrices in a pre-measured form, and simplifying provision of measured amounts of synthesis support matrix to a reaction chamber.

As used herein, the term "idle," in reference to a synthesis column, refers to columns that do not take part in a particular synthesis reaction step of a nucleic acid synthesizer. Idle synthesis columns include, but are not limited to, columns in which no synthesis occurs at all, as well as columns in which synthesis has been completed (*e.g.*, for short oligonucleotide) while other columns are actively undergoing additional synthesis steps (*e.g.*, for longer oligonucleotides).

As used herein, the term "active," in reference to a synthesis column, refers to columns that take part (or are taking part) in a particular synthesis reaction step of a nucleic acid synthesizer. Active synthesis columns include, but are not limited to, columns in which liquid reagents are being dispensed into, or columns that contain liquid reagents (*e.g.* waiting to be purged), or columns that are in the process of being purged.

As used herein, the term "O-ring" refers to a component having a circular or oval opening to accommodate and provide a seal around another component having a circular or oval external cross-section. An O-ring will generally be composed of material suitable for providing a seal, *e.g.*, a resilient air-or moisture-proof material. In some embodiments, an O-ring may be a circular opening in a larger gasket. A single gasket may contain multiple openings and thus provide multiple O-rings. In other embodiments, an O-ring may be ring-shaped, *i.e.*, it may have circular interior and exterior surfaces that are essentially concentric.

As used herein, the term "viewing window" refers to any transparent component configured to allow visual inspection of an item or material through the window. An enclosure may include a transparent portion that provides a viewing window for item within the disclosure.

Likewise, an enclosure may be made entirely of a transparent material. In such embodiments, the entire enclosure can be considered a viewing window. A "viewing window" in an enclosure that is "configured to allow visual inspection" of items in the enclosure "without opening the enclosure" refers to a viewing window in an enclosure of sufficient size, location, and transparency to allow the item to be viewed, unhindered, by the human eye. For example, where the item is one or more reagent bottles, the window is configured to allow viewing of the reagents bottles by the human eye to determine if the bottles or full or empty. A window that does not provide adequate visual inspection of each of the reagent bottles is not configured to allow visual inspection of reagents in the enclosure without opening the enclosure.

As used herein, the term "enclosure" refers to a container that separates materials contained in the enclosure from the ambient environment (*e.g.*, as in a sealed system). For example, an enclosure may be used with a reagent station to contain reagents within an interior chamber of the enclosure, and therefore separate the reagents from the ambient environment. In some embodiments, the enclosure provides an airtight or substantially airtight seal between the interior and exterior of the enclosure. The enclosure may contain one or more valves (*e.g.*, ventilation ports), doors, or other means for allowing gasses or other materials (*e.g.*, reagent bottles) to enter or leave the interior environment of the enclosure.

As used herein, the term "reaction enclosure" refers to an enclosure that separates the reaction columns or other reaction vessels (*e.g.*, microplates) from the ambient environment. For example, a chamber bowl 18 closed with a top cover 30 and sealed with a chamber seal 31 is one exemplary embodiment of a reaction enclosure. Another example of a reaction enclosure is a synthesis case, *e.g.*, as provided with a POLYPLEX synthesizer (GeneMachines, San Carlos, CA) and with the synthesizers described in WO 00/56445. In preferred embodiments, reaction enclosures can be sealed during at least one step of operation (*e.g.*, during active synthesis) and can be opened for at least one step of operation (*e.g.*, for inserting or removing reaction vessels).

As used herein, the term "top enclosure" refers to an enclosure that forms a primarily enclosed space over the top cover. In preferred embodiments, the top enclosure has four sides (*e.g.*, four top enclosure sides, *e.g.*, 98) and a top panel (*e.g.*, 97) that form a primarily enclosed space (e.g. 104) above the top cover (*e.g.*, 30) containing a plurality of valves (*e.g.*, 10) and a plurality of dispense lines (*e.g.*, 6). In some embodiments, the primarily enclosed space (*e.g.*,

104) is open to the ambient environment through a ventilation slot (*e.g.*, 100) in the top cover or the top enclosure. In certain embodiments, the top panel (*e.g.*, 99) contains an outer window (*e.g.*, 101).

Also as used herein, the combination of a "top enclosure" and "top cover" (*e.g.*, formed as one unit, or connected together) is referred to collectively as the "lid enclosure". In preferred embodiments, the "lid enclosure" (*e.g.*, 102) has six sides, with the top cover (*e.g.*, 30) serving as the "bottom", the top panel serving as the surface opposite the top cover, and the four side walls being the top enclosure sides (*e.g.*, 98). In certain embodiments, the lid enclosure is hinged so that is may be moved upward and downward.

As used herein, the term "primarily enclosed space" refers to a space having reduced contact with the ambient environment. A primarily enclosed space need not be sealed. For example, in some embodiments, a primarily enclosed space 104 of a lid enclosure of the present invention has contact with the ambient environment through a ventilation slot (*e.g.*, 100). In some embodiments, a primarily enclosed space 104 of a synthesizer base 2 has contact with the ambient environment through a ventilation slot (*e.g.*, 100)

As used herein, the term "ventilated workspace" refers to a work area that is open to the ambient environment but that is maintained under negative air pressure such that air flows into the ventilated workspace, thereby reducing or preventing the flow of fumes and emissions from the ventilated workspace into the ambient environment. One example of a ventilated workspace is a fume hood (*e.g.* a chemical fume hood). In some embodiments, the ventilated workspace that is part of an apparatus (e.g., a nucleic acid synthesizer), such that the negative air pressure is maintained over a reaction chamber to draw air away from the reaction chamber so as to prevent the air from entering the ambient environment.

As used herein, the term "synthesis" refers to the assembly of polymers from smaller units, such as monomers.

As used herein, the term "fluidic connection" refers to a continuous fluid path between components.

As used herein, the term "parallel" refers to systems or actions functioning in an essentially simultaneous, side-by-side, manner (*e.g.*, parallel synthesis or parallel synthesis system).

As used herein, the term "reaction support" refers to a structure supporting, comprising, or containing one or more reaction chambers.

As used herein, the term “rare mutation” refers to a mutation that is present in 20% or less (preferably 10% or less, more preferably 5% or less, and more preferably 1% or less) of a population of nucleic acid molecules in a sample (i.e., wherein the remaining 80% or more of the nucleic acid molecules have a wild type sequence or a different mutation in the corresponding region of the nucleic acid molecules).

As used herein, the term "distinct" in reference to signals refers to signals that can be differentiated one from another, *e.g.*, by spectral properties such as fluorescence emission wavelength, color, absorbance, mass, size, fluorescence polarization properties, charge, etc., or by capability of interaction with another moiety, such as with a chemical reagent, an enzyme, an antibody, etc.

GENERAL DESCRIPTION OF THE INVENTION

The present invention relates to detection assay development, production, usage and optimization. In particular, the present invention provides systems and methods for acquiring and analyzing biological information. The present invention also provides detection assay production with improved oligonucleotide synthesis and processing systems. The present invention further provides systems that integrate biological information collection with detection assay production that allow for rapid development of commercial products, such as analyte specific reagents (ASRs) and in *vitro* diagnostics (IVDs).

For example, the present invention provides systems and methods for the use of genetic information in the generation of assays for detecting the genetic identity of samples, the production of assays, the use of assays for gathering genetic information of individuals and populations, and the storage, analysis, and use of the obtained information, including the use of information in selecting detection assays for research use, use in panels, use as ASRs, and use in clinical diagnostics (e.g., *in vitro* diagnostics).

In some preferred embodiments, the present invention provides systems and methods for analyzing available sequence information (*e.g.*, publicly available sequence information and information obtained by the methods described herein) in the selection of informative DNA and

RNA target sequences for detections and analysis of individuals and populations. The present invention also provides systems and methods for the design and production of detection assays directed to such target sequences. The present invention further provides systems and methods for the collection, storage and analysis of data derived from detection assays.

Importantly, the present invention provides integrated systems and methods that exploit the synergies of the above systems and methods to provide comprehensive solutions, allowing for large scale and informative analysis of sequences for identifying genotype/phenotype correlations, measuring differences in gene expression, identifying allele frequencies in populations, and typing individuals and populations for important (e.g., medically relevant) sequences. For example, in some embodiments, the present invention applies data obtained from detection assays to improve the selection of target sequences, design of improved assays, and selection of assays that are suitable for use on multi-analyte panels, as ASRs, and for clinical diagnostics.

A general overview of the systems of the present invention is provided in Figure 1. The present invention provides detection assay development, production and optimization (See, section A below). For example, orders are received from customer (e.g. a target sequence is entered via a web interface), and the orders are processed (See, section A.I., "Target Sequence Selection), and Detection Assays are Designed (See Section A.III, below). The designed assays are produced (or filled from inventory) in a production facility (See, section III below). The assays that are produced are stored in inventory or shipped to customers. Preferably, each of these components are operably linked to a central data management system (e.g. running enterprise software such as Oracle), such that data and status of orders is communicated throughout the system (See, Section A.IV., below).

Detection assays are shipped to customers who use the detection assay and generate data. In certain embodiments, the data generated by the use of these detection assays is gathered, analyzed, and stored (See, section A.V, below). This information may then be integrated with the order, design, production and storage components mentioned above (See, A.VI. below). In this regard, data is continuously generated that allows, for example, an association between detection assays or targets with particular medical conditions to be established.

Gathering, analyzing, and producing detection assays while generating association data allows the clinical detection assays (e.g., ASRs and In vitro Diagnostics) to be developed and validated (See, Section, B below) through a funneling process that allows a business to focus on particularly useful assays. Assays may be incorporated in panels or databases in order to be distributed to research facilities (e.g. ASR certified), hospitals, doctors, and other customers (See, Section, C below). Employing these detection assays, or panels of assays, in a clinical setting, for example, further allows data to be collected and further associated with a patient's medical records (e.g. See, D, below). This increases the value of data that is collected and shared with the management systems of the present invention. Integrating the production systems, databases, and managements systems of the present invention allows efficient production of particular assays, as well as rapid identification of ASRs, and in vitro diagnostics. Furthermore, integration of these systems allows for accurate business pricing of various assays (See, section C, below), allowing, for example, differential pricing of ASRs and In Vitro Diagnostics.

DETAILED DESCRIPTION OF THE INVENTION

The following discussion provides a description of certain preferred illustrative embodiments of the present invention and is not intended to limit the scope of the present invention. For convenience, the discussion focuses on the application of the present invention to the detection of DNA targets, but it should be understood that the methods and systems are intended for use in the development of tools for the analysis of any nucleic acid analyte, *e.g.*, DNA or RNA. Also, for the sake of illustration, the discussion often focuses on the characterization of SNPs using INVADER assay technology. It should be understood that the methods and systems of the present invention are intended for use in detecting other biologically relevant factors using a wide variety of detection assay technologies.

As discussed above, the present invention provides systems and methods for developing detection assays for research and clinical use. The following sections describe the high throughput design, optimization, and production of detection assays in a manner that allows assays to pass from a discovery phase to use as clinical diagnostic assays. The description is provided in the following sections: A) Detection Assay Development, Production, and

Optimization; B) Development of Clinical Detection Assays; C) Distribution and Use of Detection Assays, D) Medical Records; and E) Financial Component.

A. Detection Assay Development, Production, and Optimization

The detection assay development, production, and optimization is illustrated below for hybridization-bases assays. One skilled in the art will appreciate the general applicability of various aspects of this description to other types of detection assays. The discussion of detection assay development, production, and optimization is provided in the following sections: I) Target Sequence Selection; II) Detection Assay Design; III) Detection Assay Production; IV) Data Management Systems; V) Detection Assay Use and Data Generation and Collection; and VI) Integrated Information, Design, and Production (Optimization). It will be appreciated that every step may not be required for each detection assay. For example, where a valid target sequence and assay design are already known, production and testing may be started directly. The steps may be used for original assay development and/or may be used to re-evaluate a pre-existing detection assay, whether is be for a research or a clinical detection assay. Examples of process configurations for integrating the steps (e.g., with software) are provided in Figures 1, 58, 61, and 62. As shown in Figure 1, direct clients or distributors go through an order entry process (described in detail below). Detections assays corresponding to particular oligonucleotides, primers, panels, polymorphisms (e.g., SNPs) are entered and process through an *in silico* validation process (described in detail below) and assay design software (e.g., INVADERCREATOR software). If a request corresponds to a previously validated or ordered sequences, software locates the product and proceeds with the order accordingly. Designed detection assays are then sent to a production facility for production and validation (described in detail below). Data generated by the process or from use of the detection assays and collected and stored in databases (described in detail below).

I. Target Sequence Selection

The ability to detect the presence or absence of specific target sequences in a sample underlies much of the fields of molecular diagnostics and molecular medicine. For example, tremendous effort has been expended in the development of detection assays for nucleic acid

sequence mutations that correlate to phenotypes of interest (e.g., inherited diseases). During the development of the present invention, it was found that the design of a detection assay based on a published target sequence was often not sufficient to produce viable assays. In some circumstances assays will not work at all. In others, they may work for particular individuals or populations, but fail with other individuals or populations. The present invention provides systems and methods for selecting appropriate target sequences that can be successfully targeted by detection assays.

The problem with existing methods and the solutions provided by the present invention can be illustrated by example. Many detection assays are based on the principle of nucleic acid hybridization. An oligonucleotide is designed to hybridize to a portion of the target sequence; the presence of the hybrid, or the cleavage, elongation, ligation, disassociation, or other alterations of the oligonucleotide are detected as a means for characterizing the presence or absence of the sequence of interest (e.g., a SNP). Because there is sequence heterogeneity in the population, an oligonucleotide designed to hybridize to a target sequence of one individual may not hybridize to the corresponding sequence from another individual. For example, a first individual may have a gene sequence containing a SNP that is to be detected. A second individual may have the SNP, but also may have additional sequence differences in the vicinity of the SNP that prevent the hybridization of an oligonucleotide that was designed based on the sequence of the first individual. Additionally, target sequence information obtained from a public source may contain errors (e.g., may provide the wrong sequence) or may comprise incomplete, but essential, information. For example, a given target sequence may be found in multiple locations in the genome—the intended region that the assay is designed to detect, and unintended regions that would result in false positive or otherwise misleading assay results.

The systems and methods of the present invention provide an analysis of candidate target sequences to determine if they are suitable for use in detection assays. The systems and methods of the present invention also select appropriate sequences that are likely to function in the intended detection assay. This aspect of the present invention is referred to herein as "in silico analysis," as computer analysis is conducted to analyze candidate target sequences against sequence and sequence-related information databases. In silico analysis may be performed prior

to, or in conjunction with other processes of the present invention (e.g., detection assay design and production, selection of materials for panels, ASRs, and clinical tests, etc.).

In silico analysis methods of the present invention include one or more of the following sequence analysis and processing steps: input of a candidate sequence; editing of the candidate sequence, where necessary; screening of the candidate sequence for repeat sequences; screening of the candidate sequence for research artifact sequences; identification of the candidate sequence in a sequence database; conformation of the candidate sequence in a second (or additional) sequence database; information gathering using one or more sequence information databases; problem reporting; and/or transmission of an approved target sequence for production (*e.g.*, automated production).

A. Sequence Input (Order Entry Component)

Sequences may be input for in silico analysis from any number of sources. In many embodiments, sequence information is entered into a computer. The computer need not be the same computer system that carries out in silico analysis. In some preferred embodiments, candidate target sequences may be entered into a computer linked to a communication network (e.g., a local area network, Internet or Intranet). In such embodiments, users anywhere in the world with access to a communication network may enter candidate sequences at their own locale. In some embodiments, a user interface is provided to the user over a communication network (e.g., a World Wide Web-based user interface), containing entry fields for the information required by the in silico analysis (e.g., the sequence of the candidate target sequence).

The use of a Web based user interface has several advantages. For example, by providing an entry wizard, the user interface can ensure that the user inputs the requisite amount of information in the correct format. In some embodiments, the user interface requires that the sequence information for a target sequence be of a minimum length (e.g., 20 or more, 50 or more, 100 or more nucleotides) and be in a single format (e.g., FASTA). In other embodiments, the information can be input in any format and the systems and methods of the present invention edit or alter the input information into a suitable form for in silico analysis. For example, if an input target sequence is too short, the systems and methods of the present invention search public

databases for the short sequence, and if a unique sequence is identified, convert the short sequence into a suitably long sequence by adding nucleotides on one or both of the ends of the input target sequence. Likewise, if sequence information is entered in an undesirable format or contains extraneous, non-sequence characters, the sequence can be modified to a standard format (e.g., FASTA) prior to further in silico analysis. The user interface may also collect information about the user, including, but not limited to, the name and address of the user. In some embodiments, target sequence entries are associated with a user identification code.

In certain embodiments, there is a separate component for entering large orders (e.g. entered by large companies), a separate component for entering small orders (e.g. entered by individual researchers), and a separate component for clinical orders (e.g. hospitals and clinical laboratories). In some embodiments, sequences are input directly from assay design software (e.g., the INVADERCREATOR software described below).

In preferred embodiments, each sequence is given an ID number. The ID number is linked to the target sequence being analyzed to avoid duplicate analyses. For example, if the in silico analysis determines that a target sequence corresponding to the input sequence has already been analyzed, the user is informed and given the option of by-passing in silico analysis and simply receiving previously obtained results.

The customer order component also includes one or more screens or web pages that include detection assay instrumentation data. Detection assay instrumentation data includes data describing various systems and devices, including but not limited to liquid handlers, workstations, and other automation options shown in, for example, Table 2, which are used to facilitate use of the detection assays created using the methods and systems described herein. By way of example, once a customer selects a particular type of panel format, e.g. 96 well, 385 well or 1536 well and assay configuration, he is automatically linked or presented with data of appropriate corresponding devices that are used to read the panel format which are offered for sale to the customer. In another variant, the system stores information about the type of instrumentation the customer already has in house or has previously purchased, and automatically determines and suggests the type of panel format for detection assays that the customer should buy on the customer order component, e.g. 96 well, 384 well or 1536 well. By way of further example, the customer is also provided with instrumentation pricing data,

instrument specification data, delivery data, shipping data, for various combinations of instrumentation that would suit the customer's needs. The customer order entry component can then feed data on the customer's instrumentation order (or in-house instrumentation where the customer makes a selection from an instrumentation menu presented on the web site) to the detection assay production component (including resident hardware and software components thereof) so that projections can be made as to the number and type of various detection assay starting materials that need to be purchased or stocked based upon the customers selection of instrumentation and projected usage of disposable detection assays, e.g. reagents, glass slides, plastic arrays, etc.

In yet a further embodiment, a single customer's (or a plurality of networked customers') instrumentation has a communication link to the customer order component or the detection assay production facility for exchanging data therebetween. It is appreciated that detection assay usage data is transferred from the customer's instrumentation to the detection assay production facility (or other components of the system) to help schedule and produce detection assays and order reagents and components therefore, or prompt the customer via e-mail that his stock of detection assays is nearing a predetermined number and that the customer needs to re-order detection assays. In another variant, once a threshold usage number of detection assays is determined, the customer's instrumentation automatically sends order data to the customer order component or other component of the system automatically ordering additional detection assays for one or more customers. In some embodiments, these systems are linked to a pricing component, wherein repeat customers may receive beneficial pricing for re-orders or upon reaching a total threshold volume of orders over time.

B. Web-ordering systems and methods

Users who wish to order detection assays, have detection assay designed, or gain access to databases or other information of the present invention may employ an electronic communication system (e.g., the Internet). In some embodiments, an ordering and information system of the present invention is connected to a public network to allow any user access to the information. In some embodiments, private electronic communication networks are provided. For example, where a customer or user is a repeat customer (e.g., a distributor or large diagnostic

laboratory), the full-time dedicated private connection may be provided between a computer system of the customer and a computer system of the systems of the present invention. The system may be arranged to minimize human interaction. For example, in some embodiments, inventory control software is used to monitor the number and type of detection assays in possession of the customer. A query is sent at defined intervals to determine if the customer has the appropriate number and type of detection assay, and if shortages are detected, instructions are sent to design, produce, and/or deliver additional assays to the customer. In some embodiments, the system also monitors inventory levels of the seller and in preferred embodiments, is integrated with production systems to manage production capacity and timing.

In some embodiments, a user-friendly interface is provided to facilitate selection and ordering of detection assays. Because of the hundreds of thousands of detection assays available and/or polymorphisms that the user may wish to interrogate, the user-friendly interface allows navigation through the complex set of options. For example, in some embodiments, a series of stacked databases are used to guide users to the desired products. In some embodiments, the first layer provides a display of all of the chromosomes of an organism. The user selects the chromosome or chromosomes of interest. Selection of the chromosome provides a more detailed map of the chromosome, indicating banding regions on the chromosome. Selection of the desired band leads to a map showing gene locations. One or more additional layers of detail provide base positions of polymorphisms, gene names, genome database identification tags, annotations, regions of the chromosome with pre-existing developed detection assays that are available for purchase, regions where no pre-existing developed assays exist but that are available for design and production, etc. (See, Figures 2a-f). Selecting a region, polymorphism, or detection assay takes the user to an ordering interface, where information is collected to initiate detection assay design and/or ordering. In some embodiments, a search engine is provided, where a gene name, sequence range, polymorphism or other query is entered to more immediately direct the user to the appropriate layer of information.

In certain embodiments, a user may select a PCR (or other amplification technology) or non-PCR option, depending if they want to employ amplification along with their detection assay. The PCR primer section may be employed to design such assays, taking into consideration the target and the detection assay selected by the user (see below).

In some embodiments, the ordering, design, and production systems are integrated with a finance system, where the pricing of the detection assay is determined by one or more factors: whether or not design is required, cost of goods based on the components in the detection assay, special discounts for certain customers, discounts for bulk orders, discounts for re-orders, price increases where the product is covered by intellectual property or contractual payment obligations to third parties, and price selection based on usage. For example, where detection assays are to be used for or are certified for clinical diagnostics rather than research applications, pricing is increased. In some embodiments, the pricing increase for clinical products occurs automatically. For example, in some embodiments, the systems of the present invention are linked to FDA, public publication, or other databases to determine if a product has been certified for clinical diagnostic or ASR use.

In one variant of the invention, the system and method of the present invention includes an organism-specific web order entry component. The organism-specific web order entry component comprises one or more screens and/or linked web pages that are interactively directed to present for sale one or more detection assays for a specific organism(s). By way of example, a web page or combination of web pages provides displays of the chromosomes, genes, and/or detection assays for various transgenic plants, wild type plants, wild type animals, transgenic animals, and/or genetically altered or naturally occurring microorganisms, e.g. bacteria, viruses, etc. By way of further example, one or more screens of different linked web pages permit a user to drill down into a specific genus, species and/or sub-species of an organism and/or chromosomes (or sub-parts thereof), and display the various detection assays created for the organism and/or detection assays that have been created that may be used across various organisms. The detection assays are optionally linked to specific genes or portions of chromosomes of a single organism or of multiple related or unrelated organisms.

C. In silico Processing Systems

In silico analysis utilizes one or more sequence and information databases (*e.g.*, public or private sequence databases) and software applications for processing sequence and database information (See, e.g. Figure 3). In some preferred embodiments, databases and software for in silico analysis are housed in a single location on one or more computers. Housing the databases

and processing software locally provides increased and consistent speed and access to information. In other embodiments, one or more databases and software components located on external computers are accessed over a communication network (e.g., accessed over the World Wide Web).

In preferred embodiments, databases that are maintained locally are updated regularly (*e.g.*, following each update of the web-based server, a new version is downloaded to local servers). In some preferred embodiments, databases are surveyed periodically to determine if a new version is available and, if so, one is downloaded. In some preferred embodiments, more than one copy of each database is available locally. In particularly preferred embodiments, downloaded data is parsed to extract the data, and the parsed data is configured to automatically populate the fields of one or more receiving databases (*e.g.*, an association database, a SNP database). In some embodiments, Perl scripts are used to sort data, *e.g.*, line-by-line, and to create new text files (*e.g.*, having data tagged according to the receiving field in the receiving database) for importation into the fields of a receiving database.

In some embodiments, the database analysis system comprises one or more central nodes (e.g., a computer containing a processor and computer memory) and a plurality of sub-nodes. In some embodiments, the sub-nodes house individual databases (or portions thereof) or software programs. In preferred embodiments, the central node controls the flow of information between sub-nodes, sending search requests to the sub-nodes and receiving search results from the sub-nodes. For example, in some embodiments, the central node directs data (*e.g.*, candidate target sequence) to a sub node for a database search, receives the results, and directs the information to another sub-node for additional database searching. In some preferred embodiments, the central node directs information to multiple sub nodes simultaneously (*e.g.*, for multiple concurrent database searches).

In some embodiments, in order to increase database access speed, individual databases are split among multiple (*e.g.*, two) sub-nodes. In other embodiments, databases are housed on a single node. In preferred embodiments, databases are present in multiple copies on multiple sub-nodes. In some preferred embodiments, the central node monitors database load and status on each sub-node and directs searches to the node with the greatest available capacity.

In some preferred embodiments, the central node further directs resource management software. For example, individual nodes are sent test sequences on a regular basis to ensure that they are receiving information and processing information on a desired time scale. If a sub node is found to not be functioning properly, the central node directs information to a secondary sub node containing a copy of the database. In other embodiments, sub-nodes conduct self-monitoring routines and send status reports back to the central node. For example, in some embodiments, if a search on a sub-node fails or times out, the sub-node reports this information back to the central node so that appropriate action can be taken (*e.g.*, send the search to another node and/or flag a particular sub-node for intervention). In some preferred embodiments, the central node maintains a queue of jobs submitted to each sub-node and warns human supervisors if a job fails to be completed.

In some embodiments, the central node comprises one or more workstations. In some embodiments, the sub nodes comprise two or more workstations. In other embodiments, the sub nodes comprise 5 or more workstations. In yet other embodiments, the sub nodes comprise 10 or more workstations. The present invention is not limited to a particular model or type of workstation. One skilled in the art understands that a variety of new processors of increasing speeds are regularly introduced into the market and that any suitable work station may be substituted for those described herein.

In some embodiments, in silico analysis of a candidate target sequence is completed in less than 10 seconds. In some preferred embodiments, in silico analysis of a candidate target sequence is completed in less than 2 seconds. In still more preferred embodiments, in silico analysis is completed in less than one second. In some embodiments, more than one (*e.g.*, at least 5, preferably at least 20, and even more preferably, at least 100) sequences are analyzed simultaneously using the in silico analysis system of the present invention.

1. Preliminary Sequence Screening

In some embodiments of the present invention, the first step of in silico analysis of candidate target sequences is prescreening the candidate target sequences to maximize sequence database search efficiency.

In some embodiments, candidate target sequences are searched for repeat sequences. "Repeat sequences" refers to sequences that are known to repeat multiple times in a sample (e.g., in an organism's genome). Many genomes contain large regions of repeated sequences. The presence of repeated sequences in detection assay hybridization oligonucleotides can cause the oligonucleotide to hybridize to sequences other than, and/or in addition to, the intended target. Additionally, because repeat sequences are found in multiple copies in the genome, databases searches may operate very slowly or may not proceed. In some embodiments, RepeatMasker is a perl script used in conjunction with REPBASE, which is a database of known Human repeats and is used to screen for repeat sequences. Repeat Masker screens DNA sequences for interspersed repeats and low complexity DNA sequences. Sequence information in FASTA format is input through a web-browser interface or by uploading a file. Multiple sequences may be input at once or may be contained within a file. There is no limit to the length of the query sequence or size of the batch file. Sequence comparisons in RepeatMasker are performed by the program Cross-match, an implementation of the Smith-Waterman-Gotoh algorithm developed by Phil Green. In some embodiments, RepeatMasker is run using MaskerAid (Bioinformatics 16:1040-1 [2000], available through licensing from Washington University in Saint Louis, MO), a performance enhancer for RepeatMasker. Execution profiling of native RepeatMasker showed that the vast majority of its time was spent running Cross-Match. MaskerAid allows the faster WU-BLAST search engine to substitute transparently for CrossMatch, yielding speed improvement while effectively maintaining sensitivity. MaskerAid is fundamentally a software "wrapper" around WU-BLAST that makes it appear and function very much like CrossMatch.

The output of the program is an annotation of the repeats that are present in the sequence of interest as well as a modified version of the sequence in which all the annotated repeats have been masked. The program returns three or four output files for each query. One contains the submitted sequence(s) in which all recognized interspersed or simple repeats have been masked. In the masked areas, each base is replaced with an N, so that the returned sequence is of the same length as the original. A table annotating the masked sequences as well as a table summarizing the repeat content of the query sequence is returned. Optionally, a file with alignments of the query with the matching repeats is returned as well.

Regions of low complexity, like simple tandem repeats, polypurine and AT-rich regions can lead to spurious matches in database searches. By default they are masked along with the interspersed repeats. With the option "Do not mask simple..." only interspersed repeats are masked. This may, for example, be preferred in some embodiments where the masked sequence will be analyzed by a gene prediction program. Alternatively, with the option "Only mask simple...", one can mask only the low complexity regions (*e.g.*, in some embodiments in which it is desirable to quickly locate polymorphic simple repeats in a sequence).

When checked, the repeat sequences are replaced by Xs instead of Ns. This allows one to distinguish the masked areas from possibly existing ambiguous sequences or other stretches of Ns in the original sequence. In some embodiments the use of X, N, or both may be desired for compatibility with database search engines used in the subsequent steps of the in silico analysis. In some embodiments, only the masked candidate target sequence is used in further in silico analysis. In other embodiments, both the masked and unmasked sequences are used in subsequent searches.

In certain cases, a majority or the entirety of the candidate target sequence may be masked by RepeatMasker. When this occurs, in some embodiments, a warning is sent to the user indicating that a potentially undesirable amount of the target sequence comprises repeat sequence. The user is then give the option of selecting a different target sequence or proceeding with the original sequence (or electing both options). When a decision to proceed with the sequence is selected, an unmasked version of the sequence is processed through the remaining in silico analysis steps. Where there is a portion of the original candidate target sequence that is not masked, both unmasked and masked sequences may be processed through the remaining in silico analysis steps. In some embodiments, in silico analysis is discontinued and the candidate target sequence is sent to production (Section III, below).

In some embodiments, prior to screening for repeat sequences, an analysis is performed to determine if the candidate target sequence contains undesired artifact sequences. For example, a number of sequences deposited in public databases contain vector sequence or other sequence artifacts as a result of molecular biology handling during their initial isolation and characterization. These artifact sequences often represent synthetic sequences not corresponding to a genome sequence, or inappropriately corresponding to a genome sequence other than the

intended target. Where candidate target sequences are selected that contain artifact sequences, they are more likely to fail in detection assays and are more likely to result in undesirably long search times during the remaining in silico analysis steps. For example, rather than representing a sequence that appears once in a human genome, artifact sequence may correspond to thousands of deposited database sequence that each mistakenly contain a common vector sequence.

To correct for artifact sequence, in some embodiments, the present invention employs VecScreen (available at the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health public web site). VecScreen provides a system for identifying segments of a nucleic acid sequence that may be of vector origin. VecScreen searches a query for segments that match any sequence in a specialized non-redundant vector database (UniVec). The search uses a BLAST search routine with parameters preset for optimal detection of vector contamination. Those segments of the query that match vector sequences are categorized according to the strength of the match, and their locations are displayed.

The sequence of any vector contamination should theoretically be identical to the known sequence of the vector. In practice, occasional differences are expected to arise from sequencing errors, and less frequently, from engineered variants or spontaneous mutations. The search parameters used for VecScreen are chosen to find sequence segments that are identical to known vector sequences or which deviate only slightly from the known sequence. Vector containing sequences identified are then masked.

In some embodiments, the Repeat Masker and VecScreen screening are combined into a single search. In preferred embodiments, the candidate target sequence is first screened by VecScreen, with the results then passed through Repeat Masker. Once the screening is complete, masked sequences and/or unmasked sequences are ready for database searching as described below.

2. Database Searches

In some embodiments, database searches are performed on the candidate target sequences. Databases searches are used, among other purposes, to confirm that 1) the candidate target sequence is a sequence corresponding to a known sequence, 2) the candidate target sequence corresponds to a unique sequence in the sample to be tested, and 3) the candidate target

sequence corresponds to a reliable (e.g., confirmed) sequence. The database searches are also used to gather information (allele frequencies, disease associations, variants, location in a genome, associated patents and patent applications, etc.) about the candidate target sequence. In some embodiments, the output information from the database searches is stored in a file associated with the candidate target sequence. In further embodiments, the output information is displayed to the user.

The present invention is not limited to the databases disclosed herein. Any database that provides relevant information may find use in the searches of the present invention. In some embodiments, searches are performed consecutively. In other embodiments, searches are performed concurrently. In preferred embodiments, some searches are performed consecutively and others are performed concurrently. In some embodiments, searches are performed using BLAST (Basic Local Alignment Search Tool) search mode using FASTA formatted sequences. In preferred embodiments, results from database searches are output as text files. Results are then converted to a format that is suitable for import into an Oracle database. In some embodiments, the BioJava Project is used to convert text output into an XML-like stream that is then incorporated into an Oracle database.

Other databases that are searched or used in or with various components of the invention include rat, mouse or any other organism sequence databases. It is also appreciated that the present invention can cross reference detection assays across different species of organisms. By way of example, if a customer designates a human detection assay on a customer order entry screen, the software or routines of the invention may automatically present and offer for sale on the customer's computer screen the same or similar detection assay for rats, mice or any other organism.

Descriptions of several databases that are searched in preferred embodiments of the present invention are described below.

i. SNP Databases

In preferred embodiments, candidate target sequences are first used to search several databases which catalog SNPs. The targeted databases include NCBI's dbSNP, the UK's HGBASE SNP database, the SNP Consortium database, and the Japanese Millenium Project's

SNP database. The dbSNP database serves as a central repository for both single base nucleotide substitutions and short deletion and insertion polymorphisms, and includes all the SNPs identified in the SNP Consortium effort, 10% of the Japanese SNP database and 50% of the HGBASE SNP database. The data in dbSNP is integrated with other NCBI genomic data. If a match is found in the dbSNP, the output from the search is a dbSNP accession number, which is then tied in silico to identification and characterization of genomic landscape features including known genes, predicted genes, functional location and physical location in the genome. Functional location specifies where the SNP falls within a gene or predicted gene, and details the location as exonic, promotor, intronic, 5' and 3' untranslated flanking region. The physcial location includes the base pair position of the SNP on the individual chromosome. The base pairs that make up a chromosome are counted from the p telomere to the q telomere, starting with the first base pair on the p telomere. The physical location also includes the cytoband designation that contains the SNP of interest. In some embodiments, the dbSNP search returns an accession # with an RS designation. This designation indicates that the SNP is a unique SNP identified as common between multiple studies. The RS designation is used to perform additional database mining to harvest information relating to allele frequencies, penetrance estimates and heterozyosity estimates.

ii. Gene Loci Analysis

In some embodiments, following dbSNP searches, gene loci databases (e.g., Locus Link) are searched. LocusLink provides a single query interface to curated sequence and descriptive information about genetic loci. It presents information on official nomenclature, aliases, sequence accessions, phenotypes, EC numbers, MIM numbers, UniGene clusters, homology, map locations, protein domains, and related web sites. The information output from LocusLink includes a LocusLink accession number (LocusID), an NCBI genomic contig number (NT#), a reference mRNA number (NM#), splice site variants of the reference mRNA (XM#), a reference protein number (NP#), an OMIM accession number, and a Unigene accession number (HS#).

iii. Disease Association Databases

Following the LocusLink search, the information returned is used to search disease association databases. In some embodiments, the HUGO Mutation Database Initiative, which contains a collection of links to SNP/mutation databases for specific diseases or genes, is searched.

In some embodiments, the OMIM database is searched. OMIM (Online Mendelian Inheritance in Man) is a catalog of human genes and genetic disorders developed for the World Wide Web by NCBI, the National Center for Biotechnology Information. The database contains textual information and references. Output from OMIM includes a modified accession number where multiple SNPs are associated with a genetic disorder. The number is annotated to designate the presence of multiple SNPs associated with the genetic disorder.

iv. Gene Oriented Cluster Analysis

In some embodiments, following dbSNP searches, software (*e.g.*, including but not limited to, UniGene) is used to partition search results into gene-oriented clusters. UniGene is a system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location. In addition to sequences of well-characterized genes, hundreds of thousands novel expressed sequence tag (EST) sequences are included in UniGene. Currently, sequences from human, rat, mouse, zebrafish and cow have been processed.

Unigene can be searched using either the UniGene accession number identified using LocusLink (preferred if available) or can be BLAST searched using the SNP target sequence of interest in FASTA format.

v. SNP Consortium Database

In some embodiments, masked sequences are used to search the SNP Consortium (TSC) database (available at SNP Consortium Ltd public web site). In some embodiments, SNP Consortium searches are conducted concurrently with dbSNP, LocusLink, UniGene, and OMIM searches. The SNP Consortium database includes mapping and allele frequency information. The database is searched via BLAST using the masked input target sequence. The output from

the SNP Consortium database includes a TSC accession number and a Goldenpath Contig accession number in addition to mapping and allele frequency information (if known).

vi. Genome Databases

In some embodiments, target sequences are used to search genome databases (*e.g.*, including but not limited to the Golden Path Database at University of California at Santa Cruz (UCSC) and GenBank). The GoldenPath database is searched via BLAST using the sequence in FASTA format or using the RS# obtained from dbSNP. GenBank is searched via BLAST using the masked sequence in FASTA format. In some embodiments, GoldenPath and GenBank searches are performed concurrently with TSC and dbSNP searches. In some embodiments, the searches result in the identification of the corresponding gene. Output from GenBank includes a GenBank accession number. Output from both databases includes contig accession numbers.

In some embodiments, a match to an incomplete gene is identified. In these cases, the automated system of the present invention directs the search of databases of unfinished genomic sequences (e.g., including but not limited to The High Throughput Genomic (HTG) Sequences database, a database that includes unfinished sequences from DDBJ, EMBL, and GenBank). Unfinished HTG sequences containing contigs greater than 2 kb are assigned an accession number and deposited in the HTG division. A typical HTG record might consist of all the first pass sequence data generated from a single cosmid, BAC, YAC, or P1 clone that together comprise more than 2 kb and contain one or more gaps. A single accession number is assigned to this collection of sequences and each record includes a clear indication of the status (phase 1 or 2) plus a prominent warning that the sequence data is "unfinished" and may contain errors. The accession number does not change as sequence records are updated; only the most recent version of a HTG record remains in GenBank. 'Finished' HTG sequences (phase 3) retain the same accession number, but are moved into the relevant primary GenBank division.

If a gene is identified using an unfinished sequence database, the information is transferred to the Oracle database of the present invention. If a gene is not identified, the automated system periodically (*e.g.*, weekly) searches the databases for such information.

vii. Private Databases

In some embodiments of the present invention, private databases are searched. For example, the present invention provides systems and methods for gathering, organizing, and storing sequence information (See e.g., Sections III, IV and V, below). Information obtained by the methods of the present invention may be searched during target sequence analysis to assist in the confirmation or selection of target sequences that are likely to be successful in the desired detection assay (e.g., information obtained from previously successful assays is used to select or predict successful sequences for subsequent assays on the same or similar targets using the same or similar types of detection assay).

viii. Patent Databases

In some embodiments of the present invention, patent databases are searched. In some embodiments, a search is conducted to identify patents and patent applications related to a target or probe sequence. For example, patent claims may relate to target sequences, target SNPs, probe sequences and methods of using these compositions. Searchable databases of patented sequences may be public or private. Examples of tools for searching for patented sequences include GENESEQ and The Patent Agent. GENESEQ (Derwent Information, Alexandria VA) searches for patented sequences in basic patents from 40 patent issuing authorities worldwide. GENESEQ provides a flat file (ASCII) EMBL-based format to enable integration into bioinformatics systems. The Patent Agent (DoubleTwist, Inc., Oakland, CA) uses the BLAST2N and BLAST2P algorithms to search Derwent's GENESEQ patent database and GenBank's patent division for sequence patent records matching an input (query) sequence.

3. Processing of Database Information

The collection of information obtained from the database searches is analyzed and/or stored. In some embodiments, the candidate target sequence is identified as a "high probability" target sequences and the results are reported (e.g. via the world wide web) to a user (to recommend production or use) or the target is directly sent on for production (Section III, below) or used. A high probability target sequence is one where the target sequence was confirmed to exist in one or more sequence databases, where there is no identified disagreement between the

sequence databases (*e.g.*, disagreement relating to the sequence of the target, the location of the target, or the presence of known mutations within the target region), where the target sequence represents a unique sequence in the samples that are to be assayed, and where the sequence corresponding to the target is considered reliable *(i.e.*, confirmed or completed) sequence. In some embodiments, where a report is sent to a user, the report may include results of each search, a summary of the results, a general indication that the target sequence is a high probability sequences, and/or any other detailed information identified by the searches (e.g., disease association information).

In some embodiments of the present invention, where one or more problems are identified with the candidate target sequence, a report is sent (e.g. by the internet) to a user (*e.g.*, the person who input or requested the candidate target sequence or a technician utilizing the systems and methods of the present invention) highlighting the one or more problems. Problems include the presence of repeat or artifact sequences in the candidate target sequences, multiple copies of the target sequence in the sample to be assayed (e.g., in the human genome), absence of the sequence in one or more of the databases, inconsistent results from one or more the databases (*e.g.*, inconsistency as to the sequence corresponding to the target, the location of the target within a genome, the presence or location of a mutation or SNP to be assayed, and the presence or absence of one or more additional mutations or SNPs within the target region), and/or the sequence quality (reliability) of the sequence from the databases. In some embodiments, a reliability score is generated based on the presence or absence of one or more of the above potential problems. The reliability score may be sent to the user, or may be used as a signal to cause a further action, such as to begin production and/or to cancel the candidate target sequence.

In some embodiments, the user is given the option to select another target sequence or to proceed with the present target sequence (*e.g.*, to proceed to production). In some embodiments, when problems are identified, the systems of the present invention automatically select and test additional candidate target sequences based on the original requested candidate target sequence (*e.g.*, select neighboring sequences and/or remove problem portions of the sequence). If more reliable sequences are identified, these suggested alternate target sequences are reported to the user.

An overview of in silico analysis in some preferred embodiments of the present invention is shown in Figure 3. The three top boxes represent exemplary sources of target sequences: research & development (e.g., direct input by research personnel) (20), Web interface (sequence input through a communication network) (21), and system administrators (*e.g.*, to test the systems and methods of the present invention) (22). The target sequences are then analyzed by a screening component (23) that masks repeat and artifact sequences. If sequences are suitable for further analysis, they are passed to a series of databases. In the example shown in Figure 3, the sequences are simultaneously sent to dbSNP (24), GoldenPath (25), and SNP Consortium (26) databases. If a dbSNP accession number is available, dbSNP data (27) is collected and stored and the dbSNP accession number is used to search the Unigene database (29). The dbSNP accession number may also be used to search the OMIM database (28) (which may also be searched after any other database search). If a dbSNP accession is not identified, the target sequence information is passed to the Unigene database (29). If a Unigene identification is found, Unigene data (30) is collected and stored.

The target sequence information sent to the GoldenPath database (25) is used to identify the base pair position of the SNP on the current GoldenPath assembly of the genome and to check the reliability status of the sequence. If the sequence is considered "finished" sequence, GoldenPath data is collected and stored. If the sequence is not finished, the GenBank database (31) is searched to identify a GenBank contig identification number and to determine if the contig is considered "finished." If the contig is finished, data is collected and stored. If the contig is not considered finished, a request for additional sequence data is placed with the group responsible with finishing the sequence of the region (32). If sequence data is available, data from the finishing group is collected and stored. The base pair position of the SNP generates the next level of in silico analysis to generate the genomic landscape information for each SNP resulting in a detailed in silico annotation of the SNP. The annotation is extended to include the full target sequence information. Target sequences which fall within a known gene region defined as "genic" to include 10 kilobases of sequence 5' and 3' of the beginning and end of transcription, then a second round of in silico annotation charaterizes this genic region as well.

The target sequence information sent to the SNP Consortium database (26) is used to identify a TSC identification number and TSC data, if available, is collected and stored. In some

embodiments, one or more database accession numbers (e.g., LocusLink accession number) are provided during the original target sequence input or at any time thereafter, and said accession numbers are used to direct searches in the corresponding database (e.g., LocusLink database) or other databases. To the extent that databases searches are conducted solely to obtain an accession number for use in searching other databases, pre-entry of the accession number reduced the time required for in silico analysis. All of the collected data is stored in a database and used to generate reports and/or reliability scores for use in determining whether production of an assay directed at the target sequence should proceed. In some embodiments, if production is to proceed, information from the in silico analysis, and design analysis (Section II, below) is sent to a production facility. The flow of information from sequence input to production in some embodiments of the present invention is shown Figure 4.

4. Comprehensive Approach to Whole Genome SNP Analysis and Bioinformatics

As a result of Human Genome Project (HGP), over 35 gigabytes of data is currently available in a large number of public databases, and there is now the potential to quickly and accurately describe the relationship between individual genotype and disease phenotype as never before by analyzing sequence variation. The International SNP Map Working group has constructed a map of 1.4 million candidate SNPs and estimates that two individuals differ at a rate of 1 nucleotide every 1.3 kb (2001). NCBI's dbSNP catalogs over 3 million individual and 1.8 million consensus sequence variations, Japan's SNP db catalogs 117 thousand sequence variations, and HGBASE SNP db catalogs over 65 thousand SNPs. Kruglyak and Nickerson (2001) hypothesized that this collection of sequence variations represents only 11% to 12% of the total human polymorphic nucleotide variation. Therefore, the challenge of discovery is shifting away from discovery to the planning, development, and implementation of clinically relevant assays and studies to provide a synergy between sequence data and large volumes of genotype/phenotype data with effective utilization of a platform of statistical analysis to define disease associations. Additionally, developing and implementing strategies to convert genomic sequence data of varying quality and completeness into biologically meaningful information will be a key to capitalizing on this wealth of information. While the resources available from the

HGP make it possible to pursue this strategy of "targeted genomics," the efficient integration and interpretation of public databases is a major task and becomes one of the critical features of the post-sequencing era. Coupling the computational analysis of publicly available sequence data with clinical studies is crucial.

Through the *in silico* sequence analysis pipeline of the present invention, it is possible to mine the data generated by the Human Genome Project and to harvest information to annotate the genomic landscape surrounding each SNP (See Figure 5). The detailed annotation integrates Medline and OMIM data and is used to populate panels of Third Wave Technologies INVADER assays or other detection assays targeted to address specific questions related to disease gene discovery, disease susceptibility, diagnosis and treatment. The panels are designed to map genes, to characterize novel mutations, to create disease-specific gene expression snapshots, to detect clinically relevant mutations, and to facilitate and direct clinical trials of novel treatments for disease. Allele frequency information is generated for each SNP and provides integration between each SNP and the published genetic and physical maps, as well as test algorithms for the prediction of the functional impact of amino acid changes in cSNPs.

Furthermore, the in silico analysis systems and methods described above allow the rapid development of products such as Analyte-Specific Reagents and In-vitro Diagnostics. Since the in silico analysis integrates sequence and expression data with literature and clinical data (e.g. data is fed back into the data management systems of the present invention) the product development funnel (See, section B.IV) if further promoted (See, Figure 5).

5. RNA Target Sequence Selection in Gene Expression Analysis

Unlike SNP assays wherein there are only two nucleotide locations to design for (sense and antisense strands at the position of the variation), gene expression (GE) assays can be designed to numerous sites (*e.g.*, from about 100 to several 1000 different sites) in a particular mRNA sequence. Further complicating the design process is determining whether there is any homology between the RNA sequence of interest and any others that may be or are likely to be present in the sample. Homologies between target RNA and non-target RNAs occur not only in closely related gene families, but also when RNAs such as mRNAs have several alternative splice configurations. In some embodiments, the assay is intended to detect all or most members

of a set of homologous DNAs or RNAs. In other embodiments, an assay is intended to detect a particular nucleic acid and to avoid detecting any similar or related sequences present in a sample. If significant homologies exist, sequence alignments performed before the assay is designed can identify sequences unique to a particular target from sequences that are shared. SNP variations that occur in the mRNA also need to be considered, as their position in the target region can affect assay performance, and location at or near the probe cleavage site may preclude detection of that particular variant. In some embodiments, this is a preferred effect; in some embodiments it is desirable to avoid this effect.

Strategies for designing INVADER assays for detection of RNA include targeting: *i*) splice sites, *ii*) accessible sites, and *iii*) discrimination sites. The type of bioinformatic analysis performed on a given RNA target sequence depends on the type of design strategy being used for developing the assay.

Bioinformatic analysis in mRNA target sequence selection may include mapping of splice sites within the mRNA sequence, identification of any variations in the mRNA sequence (e.g. single-base changes, insertions, deletions), identification and alignment of splice variants, identification and alignment of closely related genes, homology to and alignment of the corresponding gene in other species, and location of accessible sites (unstructured regions of RNA) via *in silico* analysis. In some embodiments, sequences are obtained from and compared to information from a public database. In other embodiments, sequences are obtained from a private database and compared to information from a private and/or public database In other embodiments, relevant sequences are collected into a local database for rapid retrieval.

In some embodiments, a fully integrated bioinformatic module includes complete analysis of the RNA target sequence prior to assay design, independent of how the assay will be designed. For example, in some embodiments, the user enters a GenBank NM_ accession number and the module retrieves the sequence, compares it to an mRNA sequence database (*e.g.*, using BLAST) to retrieve sequences having a percent identity selected by the user (*e.g.*, a minimum identity of 90%), aligns the target sequence with the retrieved sequences, and then uses subroutines to output positions where there is discrimination (e.g., 2 adjacent nucleotides) compared to the collection of retrieved sequences. In some embodiments, additional subroutines comprise locating completely homologous regions of sequence relative to the collection of

retrieved sequences for the design of inclusive assays (*e.g.*, assays designed to detect all members of the collection). In other embodiments, subroutines are implemented that retrieve all known alternatively spliced variants, align them, and output splice junctions and included exons for the design of assays that either inclusively or exclusively detect these variants.

In some embodiments, a subroutine performs a BLAST comparison of the mRNA sequence from one species against other databases for other species. In some embodiments, the output of the bioinformatics module comprises identification of splice sites for each RNA.

In some embodiments, homologies are identified and used to design inclusive (*e.g.*, interspecies) assays For example, single assays can detect human and rat CYP1A1, or mouse and rat GAPDH, etc. Interspecies assays have the benefits of making product development more efficient and less expensive, since two or more assays are developed, packaged, and inventoried for the time and price of one. In some embodiments, homologies are identified and used to design exclusive assays (*e.g.*, assays that will not cross-react between species).

In some embodiments, the output of a bioinformatics module is exported to an INVADERCREATOR module. In some embodiments the information is manually entered into the INVADERCREATOR software, while in other embodiments it is read in, *e.g.*, via a batch file. In preferred embodiments, batch files comprise numerical locations for sequences selected as targets for assay design. In other embodiments, other relevant information for assay design such as full gene names, gene name abbreviations, locations of SNP variations are included in the batch files for direct import into INVADERCREATOR software.

In some embodiments, the user selects a design method after reviewing the contents of the bioinformatics output file. In other embodiments, a pre-selected or default design method based on the content of the output file is automatically selected. In some embodiments, *e.g.*, for design of an exclusive assay, the bioinformatics module exports data having particular information regarding homologous sequences found, *e.g.*, a threshold percentage identity value, and this output information directs the INVADERCREATOR module to default to a discrimination sites design method. In some preferred embodiments, information is cross-referenced in the INVADERLOCATOR software.

In some embodiments, output from an INVADERCREATOR analysis is fed back into the bioinformatics module for further analysis. In some embodiments, the bioinformatics

module verifies a design feature, *e.g.*, verifies that the final design selection(s) have the intended inclusivity or exclusivity. In other embodiments, a target selected based on one set of criteria (*e.g.*, exclusivity within the RNAs of a single species) is compared to a database using different criteria (*e.g.*, cross-species homologies). In preferred embodiments, the output of the second analysis in the bioinformatics module is returned to the INVADERCREATOR module and the user is offered the option of altering an aspect of the assay design. In other preferred embodiments, alteration or refinement of the assay design is an automated step based on the output from the informatics analysis.

In some embodiments, inventoried assay sequences are reviewed against newly updated databases. In preferred embodiments, users are notified of new information (e.g., via INVADERLOCATOR software) related to previously characterized target sequences, such as newly identified SNPs or splice variants.

II. Detection Assay Design

There are a wide variety of detection technologies available for determining the sequence of a target nucleic acid at one or more locations. For example, there are numerous technologies available for detecting the presence or absence of SNPs. Many of these techniques require the use of an oligonucleotide to hybridize to the target. Depending on the assay used, the oligonucleotide is then cleaved, elongated, ligated, disassociated, or otherwise altered, wherein its behavior in the assay is monitored as a means for characterizing the sequence of the target nucleic acid. A number of these technologies are described in detail, in Section V, below.

The present invention provides systems and methods for the design of oligonucleotides for use in detection assays. In particular, the present invention provides systems and methods for the design of oligonucleotides that successfully hybridize to appropriate regions of target nucleic acids (e.g., regions of target nucleic acids that do not contain secondary structure) under the desired reaction conditions (e.g., temperature, buffer conditions, etc.) for the detection assay. The systems and methods also allow for the design of multiple different oligonucleotides (e.g., oligonucleotides that hybridize to different portions of a target nucleic acid or that hybridize to two or more different target nucleic acids) that all function in the detection assay under the same or substantially the same reaction conditions. These systems and methods may also be used to

design control samples that work under the experimental reaction conditions. The present invention also provides methods for designing sequences for amplifying the target sequence to be detected (e.g. designing PCR primers for multiplex PCR).

While the systems and methods of the present invention are not limited to any particular detection assay, the following description illustrates the invention when used in conjunction with the INVADER assay (Third Wave Technologies, Madison WI; See e.g. U.S. Patent Nos. 5,846,717; 6,090,543; 6,001,567; 5,985,557; 5,994,069, 6,214,545, 6,210,880, and 6,194,880; Lyamichev et al., Nat. Biotech., 17:292 (1999), Hall et al., PNAS, USA, 97:8272 (2000), Agarwal et al., Diagn. Mol. Pathol. 9:158 [2000], Cooksey et al., Antimicrob. Agents Chemother. 44:1296 [2000], Griffin and Smith, Trends Biotechnol., 18:77 [2000], Griffin and Smith, Analytical Chemistry 72:3298 [2000], Hessner et al., Clin. Chem. 46:1051 [2000], Ledford et al., J. Molec. Diagnostics 2,:97 [2000], Lyamichev et al., Biochemistry 39:9523 [2000], Mein et al., Genome Res., 10:330 [2000], Neri et al., Advances in Nucleic Acid and Protein Analysis 3826:117 [2000], Fors et al., Pharmacogenomics 1:219 [2000], Griffin et al., Proc. Natl. Acad. Sci. USA 96:6301 [1999], Kwiatkowski et al., Mol. Diagn. 4:353 [1999], and Ryan et al., Mol. Diagn. 4:135 [1999], Ma et al., J. Biol. Chem., 275:24693 [2000], Reynaldo et al., J. Mol. Biol., 297:511 [2000], and Kaiser et al., J. Biol. Chem., 274:21387 [1999]; and PCT publications WO97/27214, WO98/42873, and WO98/50403, each of which is herein incorporated by reference in their entirety for all purposes) to illustrate preferred features of the present invention) to detect a SNP or other sequence of interest. The INVADER assay provides ease-of-use and sensitivity levels that, when used in conjunction with the systems and methods of the present invention, find use in detection panels, ASRs, and clinical diagnostics. One skilled in the art will appreciate that specific and general features of this illustrative example are generally applicable to other detection assays.

A. INVADER Assay

The INVADER assay provides means for forming a nucleic acid cleavage structure that is dependent upon the presence of a target nucleic acid and cleaving the nucleic acid cleavage structure so as to release distinctive cleavage products (See, Figure 6). 5' nuclease activity, for example, is used to cleave the target-dependent cleavage structure and the resulting cleavage

products are indicative of the presence of specific target nucleic acid sequences in the sample. When two strands of nucleic acid, or oligonucleotides, both hybridize to a target nucleic acid strand such that they form an overlapping invasive cleavage structure, as described below, invasive cleavage can occur. Through the interaction of a cleavage agent (*e.g.*, a 5' nuclease) and the upstream oligonucleotide, the cleavage agent can be made to cleave the downstream oligonucleotide at an internal site in such a way that a distinctive fragment is produced.

The INVADER assay provides detections assays in which the target nucleic acid is reused or recycled during multiple rounds of hybridization with oligonucleotide probes and cleavage of the probes without the need to use temperature cycling (*i.e.*, for periodic denaturation of target nucleic acid strands) or nucleic acid synthesis (*i.e.*, for the polymerization-based displacement of target or probe nucleic acid strands). When a cleavage reaction is run under conditions in which the probes are continuously replaced on the target strand (*e.g.* through probe-probe displacement or through an equilibrium between probe/target association and disassociation, or through a combination comprising these mechanisms, (Reynaldo, *et al.*, J. Mol. Biol. 97: 511-520 [2000]), multiple probes can hybridize to the same target, allowing multiple cleavages, and the generation of multiple cleavage products.

The INVADER assay, as well as other assays, may also employ degenerate oligonucleotides (e.g. degenerate INVADER and probe oligonucleotides). For example, standard INVADER oligonucleotides and probes may be randomly changed at one more positions such that a set of degenerate INVADER and/or probe oligonucleotides are produced. Degenerate sets of INVADER and probe oligonucleotides are particularly useful for use in conjunction with target sequences that tend to be heavily mutated (e.g. HIV-1 pol gene). Using such degenerate sets of INVADER and probe oligonucleotides allows the presence of target sequences at a particular location to be detected even if the surrounding sequence no longer represent the wild type or expected sequence.

The INVADER assay technology may be used to quantitate mRNA (e.g. without target amplification). Low variability (3-10 % coefficient of variation) provides accurate quantitation of less than two-fold changes in mRNA levels. A biplex FRET-based detection format enables simultaneous quantitation of expression from two genes within the same sample. One of these genes can be an invariant housekeeping gene that is used as the internal standard. Normalizing

the signals from the gene of interest with the internal standard provides accurate results and obviates the need for replicate samples. A simple and rapid cell lysate sample preparation method can be used with the mRNA INVADER Assay. The combined features of biplex detection and easy sample preparation make this assay readily adaptable for use in high-throughput applications.

In certain embodiments, the INVADER assay (and other detection assays such as TAQMAN) employ an E-TAG label (e.g. as part of the INVADER oligonucleotide, probe oligonucleotide, or the FRET oligonucleotide). E-TAG labeling is particularly useful in muliplex analysis. E-TAG labeling does not require surface immobilization of affinity agents. E-TAG type labeling is described in U.S. Pat. 5,858,188; 5,883,211; 5,935,401; 6,007,690; 6,043,036; 6,054,034; 6,056,860; 6,074,827; 6,093,296; 6,103,199; 6,103,537; 6,176,962; and 6,284,113, all of which are herein incorporated by reference.

B. Oligonucleotide Design for the INVADER assay

The application of the INVADER assay is not limited to any particular type of nucleic acid or nucleic acid variations. In some embodiments, oligonucleotides for an INVADER assay are designed to detect a particular SNP. In other embodiments, the oligonucleotides for an assay may be designed to determine the presence or absence of a particular nucleic acid in a sample, *e.g.*, a nucleic acid suspected to be present as a consequence of, for example, transfection, transformation or infection of the source of the sample. In yet other embodiments, the oligonucleotides of an INVADER assay may be designed to provide quantitative information about a particular DNA or RNA sequence.

In some embodiments where an oligonucleotide is designed for use in the INVADER assay, the sequence(s) of interest are entered into the INVADERCREATOR program (Third Wave Technologies, Madison, WI). One skilled in the art will appreciate that applicability of aspects of this design system for use in other detection assays. As described above, sequences may be input for analysis from any number of sources, either directly into the computer hosting the INVADERCREATOR program, or via a remote computer linked through a communication network (*e.g.*, a LAN, Intranet or Internet network). For detection of double-stranded nucleic acid, *e.g.*, a gene, the program designs probes for both strands, *e.g.*, the sense and antisense

strands. Selection of a particular strand for detection is generally based upon factors that include the ease of synthesis, minimization of secondary structure formation, manufacturability and INVADERCREATOR penalty scores, which have been established by studying probe design performance in the INVADER assay. In some embodiments, the user chooses the strand for sequences to be designed for. In other embodiments, the software automatically selects the strand. By incorporating thermodynamic parameters for optimum probe cycling and signal generation (*e.g.*, Allawi and SantaLucia, Biochemistry, 36:10581 [1997] for DNA duplexes, Sugimoto, *et al.*, Biochemistry 34, 11211 [1995] for RNA/DNA hybrids, or Xia, *et al.*, Biochemistry 37:14719 [1998], for RNA duplexes), oligonucleotide probes may be designed to operate at a pre-selected assay temperature (*e.g.*, 63°C). Based on these criteria, a final probe set (*e.g.*, primary probes for 2 alleles and an INVADER oligonucleotide for a SNP detection assay, or primary probe, a stacker oligonucleotide, an INVADER oligonucleotide and an ARRESTOR oligonucleotide for an RNA detection assay) is selected.

In some embodiments, the INVADERCREATOR system is a web-based program with secure site access that contains a link to BLAST (available at the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health website) and that can be linked to RNAstructure (Mathews *et al.*, RNA 5:1458 [1999]), a software program that utilizes mfold (Zuker, Science, 244:48 [1989]). RNAstructure can test the proposed oligonucleotide designs generated by INVADERCREATOR for potential uni- and bimolecular complex formation. INVADERCREATOR is open database connectivity (ODBC)-compliant and uses the Oracle database for export/integration. The INVADERCREATOR system is configured with ORACLE to work well with UNIX systems, as most genome centers are UNIX-based.

In some embodiments, the INVADERCREATOR analysis is provided on a separate server (*e.g.*, a Sun server) so it can handle analysis of large batch jobs. For example, a customer can submit up to 2,000 SNP sequences in one email. The server passes the batch of sequences on to the INVADERCREATOR software, and, when initiated, the program designs detection assay oligonucleotide sets. In some embodiments, probe set designs are returned to the user within 24 hours of receipt of the sequences.

Each INVADER reaction includes at least two target sequence-specific, unlabeled oligonucleotides for the primary reaction: an upstream INVADER oligonucleotide and a downstream Probe oligonucleotide. The INVADER oligonucleotide is generally designed to bind stably at the reaction temperature, while the probe is designed to freely associate and disassociate with the target strand, with cleavage occurring only when an uncut probe hybridizes adjacent to an overlapping INVADER oligonucleotide. In some embodiments, the probe includes a 5' flap or "arm" that is not complementary to the target, and this flap is released from the probe when cleavage occurs. In some embodiments, the released flap participates as an INVADER oligonucleotide in a secondary reaction. In some embodiments, the INVADER reaction may comprise additional oligonucleotides, such as stacker or ARRESTOR oligonucleotides. In some embodiments, the designed oligonucleotides are submitted as a synthesis order, such that manufacture of each oligonucleotide is initiated at order submission, are tracked through the modules of synthesis and the manufactured set of oligonucleotides are collected into a finished assay product or kit. In other embodiments, the oligonucleotide designs are checked against an inventory of existing oligonucleotides to determine if any of the oligonucleotides of the assay have been previously synthesized ("pre-synthesized" oligonucleotides) and stored. In some embodiments, one or more pre-synthesized oligonucleotides are taken from inventory oligonucleotides and included with newly designed and synthesized oligonucleotides in the finished assay or kit. In other embodiments, new assays or kits are assembled entirely from pre-synthesized oligonucleotides taken from an inventory of oligonucleotides.

In some embodiments, of an INVADERCREATOR program, the program is configured to design oligonucleotides for an assay of a single particular type or purpose (*e.g.*, for SNP detection or RNA quantitation). In other embodiments, an INVADERCREATOR program is configured to allow a user to select, *e.g.*, through a button, check box or menu, from a variety of assay types or purposes. The following discussion provides several examples of how a user interface for an INVADERCREATOR program may be configured. Examples of user interfaces are presented in Figures 12 through 14. Figure 12 provides screens images showing one example of using an INVADERCREATOR program to designs an assay for the detection of a SNP (a SNP INVADERCREATOR, or SIC program module). Figure 13 provides a selection of screen

images showing one example of using an INVADERCREATOR program to design an assay for the detection of an RNA target (an RNA INVADERCREATOR, or RIC program module). Figure 14 provides a selection of screen images showing one example of using an INVADERCREATOR program to design an assay for the detection of a transgene (a Transgene INVADERCREATOR, or TIC program module).

In some embodiments, screens provide optional selection of any number of modifications (e.g., arms, dyes, detectable moieties) for detection or further manipulation. In some embodiments, an INVADERCREATOR module may be customized for a particular assay, or for the needs of a particular user or customer. For example, if a customer has a particular detection platform requiring that the cleavage products comprise moiety X, an INVADERCREATOR module can be configured such that all assays designed by or for customer X are automatically configured to comprise moiety X, in accordance with the customer's requirements. In some embodiments, a pre-designated design feature cannot be altered by an operator creating a new probe design using the customized INVADERCREATOR module. In other embodiments, a pre-designated design feature may be presented to an operator as a default condition of the design that may be overridden during probe design (*e.g.*, by selecting an alternative configuration through one or more data entry screens).

In one embodiment of an INVADERCREATOR program, the user initiates oligonucleotide design by opening a work screen (*e.g.*, Figures 12A, 13A or 14A), *e.g.*, by clicking on an icon on a desktop display of a computer (*e.g.*, a Windows desktop). In some embodiments, the user enters information related to the assay, such as project code, company name, assay name, etc. In some embodiments, the used indicates what species the nucleic acid sequence is from. In some embodiments, the user selects the INVADERCREATOR program module to be used (*e.g.*, SIC, RIC, TIC, etc.), *e.g.*, by clicking a button on the screen. The user enters information related to the target sequence for which an assay is to be designed. In some embodiments, the user enters a target sequence (*e.g.*, Figures 12B, 13C, or 14B). In other embodiments, the user enters a code or number that causes retrieval of a sequence from a database. In still other embodiments, additional information may be provided, such as the user's name, an identifying number associated with a target sequence, and/or an order number. In preferred embodiments, the user indicates (*e.g.* via a check box or drop down menu) that the

target nucleic acid is DNA or RNA. In other preferred embodiments, the user indicates the species from which the nucleic acid is derived. In particularly preferred embodiments, the user indicates whether the design is for monoplex (*i.e.*, one target sequence or allele per reaction) or multiplex (*i.e.*, multiple target sequences or alleles per reaction) detection. When the requisite choices and entries are complete, the user starts the analysis process. In one embodiment, the user clicks a "Design It" button to continue.

In some embodiments, the software validates the field entries before proceeding. In some embodiments, the software verifies that any required fields are completed with the appropriate type of information. In other embodiments, the software verifies that the input sequence meets selected requirements (*e.g.*, minimum or maximum length, DNA or RNA content). If entries in any field are not found to be valid, an error message or dialog box may appear. In preferred embodiments, the error message indicates which field is incomplete and/or incorrect. Once a sequence entry is verified, the software proceeds with the assay design.

In some embodiments, the information supplied in the order entry fields specifies what type of design will be created. In preferred embodiments, the target sequence and multiplex check box specify which type of design to create. Design options include but are not limited to SNP assay, Multiplexed SNP assay (*e.g.*, wherein probe sets for different alleles are to be combined in a single reaction), Multiple SNP assay (*e.g.*, wherein an input sequence has multiple sites of variation for which probe sets are to be designed), and Multiple Probe Arm assays.

In some embodiments, the INVADERCREATOR software is started via a Web Order Entry (WebOE) process (*i.e.*, through an Intra/Internet browser interface) and these parameters are transferred from the WebOE via applet <param> tags, rather than entered through menus or check boxes.

In the case of Multiple SNP Designs, the user chooses two or more designs to work with. In some embodiments, this selection opens a new screen view (*e.g.*, a Multiple SNP Design Selection view Figure 8). In some embodiments, the software creates designs for each locus specified in the target sequence, scoring each, and presents them to the user in this screen view. The user can then choose any two designs to work with. In some embodiments, the user chooses a first and second design (*e.g.*, via a menu or buttons) and clicks a "Design It" button to continue.

To select a probe sequence that will perform optimally at a pre-selected reaction temperature, the melting temperature (T_m) of the SNP to be detected is calculated using the nearest-neighbor model and published parameters for DNA duplex formation (Allawi and SantaLucia, Biochemistry, 36:10581 [1997], SantaLucia, Proc Natl Acad Sci U S A., 95(4):1460 [1998]). In embodiments wherein the target strand is RNA, parameters appropriate for RNA/DNA heteroduplex formation may be used. Because the assay's salt concentrations are often different than the solution conditions in which the nearest-neighbor parameters were obtained (1M NaCl and no divalent metals), an adjustment should be made to the value provided for the salt concentration within the melting temperature calculations. This adjustment is termed a 'salt correction' SantaLucia, Proc Natl Acad Sci U S A., 95(4):1460 [1998]. Similarly, the presence and concentration of the enzyme influence optimal reaction temperature. One way of compensating for these additional factors is to further vary the salt value in the Tm calculations. As used herein, the term "salt correction" refers to a variation made in the value provided for a salt concentration for the purpose of reflecting the effect on a T_m calculation for a nucleic acid duplex of a both an alternative salt effect and a non-salt parameter or condition affecting said duplex. Variation of the values provided for the strand concentrations will also affect the outcome of these calculations. By using a value of 0.5 M NaCl (SantaLucia, Proc Natl Acad Sci U S A, 95:1460 [1998]) and strand concentrations of about 1 M of the probe and 1 fM target, the algorithm used for calculating probe-target melting temperature has been adapted for use in predicting optimal INVADER assay reaction temperatures. For one set of 30 probes, the average deviation between optimal assay temperatures calculated by this method and those experimentally determined is about 1.5 °C.

The length of the target-complementary region of a probe (*e.g.*, the probe to a given SNP) is defined by the temperature selected for running the reaction (*e.g.*, 63°C). Starting from the target base that is paired to the probe nucleotide 5' of the intended cleavage site (*e.g.*, the position of the variant nucleotide on the target DNA)), and adding on the 3' end, an iterative procedure is used by which the length of the target-binding region of the probe is increased by one base pair at a time until a calculated optimal reaction temperature (T_m plus salt correction to compensate for enzyme effect) matching the desired reaction temperature is reached. For INVADER assays detecting DNA targets, the non-complementary arm of the probe is preferably selected to allow

the secondary reaction to cycle at the same reaction temperature. The entire probe oligonucleotide is screened using programs such as mfold (Zuker, Science, 244: 48 [1989]) or Oligo 5.0 (Rychlik and Rhoads, Nucleic Acids Res, 17: 8543 [1989]) for the possible formation of dimer complexes or secondary structures that could interfere with the reaction. The same principles are also followed for INVADER oligonucleotide design. Briefly, starting from the position N on the target DNA, additional residues complementary to the target DNA starting from residue N-1 are then added in the 5' direction until the stability of the INVADER oligonucleotide-target hybrid exceeds that of the probe (and therefore the planned assay reaction temperature), generally by 15-20 °C. The 3' end of the INVADER oligonucleotide is designed to have a nucleotide not complementary to either allele suspected of being contained in the sample to be tested. The mismatch does not adversely affect cleavage (Lyamichev *et al.*, Nature Biotechnology, 17: 292 [1999]), and it can enhance probe cycling, presumably by minimizing coaxial stabilization effects between the two probes.

It is one aspect of the assay design that all of the probe sequences may be selected to allow the primary and secondary reactions to occur at the same optimal temperature, so that the reaction steps can run simultaneously. In an alternative embodiment, the probes may be designed to operate at different optimal temperatures, so that the reaction steps are not simultaneously at their temperature optima.

In some embodiments, the software provides the user an opportunity to change various aspects of the design including but not limited to: probe, target and INVADER oligonucleotide temperature optima and concentrations; blocking groups; probe arms; dyes, capping groups and other adducts; individual bases of the probes and targets (*e.g.*, adding or deleting bases from the end of targets and/or probes, or changing internal bases in the INVADER and/or probe and/or target oligonucleotides). In some embodiments, changes are made by selection from a menu. In other embodiments, changes are entered into text or dialog boxes. In preferred embodiments, this option opens a new screen (*e.g.*, a Designer Worksheet view, Figure 9).

In some embodiments, the software provides a scoring system to indicate the quality (*e.g.*, the likelihood of performance) of the assay designs. In one embodiment, the scoring system includes a starting score of points (*e.g.*, 100 points) wherein the starting score is indicative of an ideal design, and wherein design features known or suspected to have an adverse

affect on assay performance are assigned penalty values. Penalty values may vary depending on assay parameters other than the sequences, including but not limited to the type of assay for which the design is intended (*e.g.*, DNA, RNA, monoplex, multiplex) and the temperature at which the assay reaction will be performed. The following example provides illustrative scoring criteria for use with some embodiments of the INVADER assay based on an intelligence defined by experimentation.

Examples of design features in assays for DNA detection that may incur score penalties (*e.g.*, SIC and TIC module penalties) include but are not limited to the following [penalty values are indicated in brackets; if there are 2 numbers, the first number is for lower temperature assays (*e.g.*, 62-64 °C), second is for higher temperature assays (*e.g.*, 65-66 °C)]:

1. [20] 3′ four bases of the INVADER oligonucleotide resembles the probe arm, for example:

	ARM SEQUENCE	PENALTY AWARDED IF INVADER ENDS IN:
Arm 1:	CGCGCCGAGG	5′ ......GAGGX or 5′.......GAGGXX
Arm 2:	ATGACGTGGCAGAC	5′.......AGACX or 5′........AGACXX
Arm 3:	ACGGACGCGGAG	5′.......GGAGX or 5′ ....GGAGXX
Arm 4:	TCCGCGCGTCC	5′.......GTCCX or 5′........GTCCXX

2. [100] 3′ five bases of the INVADER oligonucleotide resembles the probe arm. for example:

	ARM SEQUENCE	PENALTY AWARDED IF INVADER ENDS IN:
Arm 1:	CGCGCCGAGG	5′ ......CGAGGX or 5′........CGAGGXX
Arm 2:	ATGACGTGGCAGAC	5′.......CAGACX or 5′........CAGACXX
Arm 3:	ACGGACGCGGAG	5′.......CGGAGX or 5′ ....CGGAGXX
Arm 4:	TCCGCGCGTCC	5′.......CGTCCX or 5′........CGTCCXX

3. [70] probe has a 5-base stretch containing the polymorphism
4. [60] probe has a 5-base stretch adjacent to the polymorphism
5. [15] probe has a 4-base stretch of Gs containing the polymorphism
6. [50] probe has a 5-base stretch of Gs – penalty added anytime it is infringed
7. [40] INVADER oligonucleotide 6-base stretch is of Gs - additional penalty

8. [90] two or three base sequence repeats at least four times starting in the region +1 to +4 of the probe.

9. [100] degenerate base occurs in the probe four bases from either end.

10. [100] probe hybridizing region is short ≤ 12 bases regardless of assay temperature.

11. [40] probe hybridizing region is long (≥ 26 bases).

12. [5] hybridizing region length exceeding 26 - per base additional penalty

13, [80] insertion/deletion design with poor discrimination in first 3 bases after probe arm

14. [100] calculated INVADER oligonucleotide Tm < 7.5C of probe target Tm

15. [100] a probe has a calculated Tm 2C less than its target Tm

Tie Breaker rules for SIC module:

1. If calculated probes Tms differ by more than 2.0C, then pick other strand for design.
2. If target of one strand 8 bases longer than that of other strand, then pick shorter strand.

Examples of design features in assays for RNA detection (*e.g.*, RIC module penalties) that may incur score penalties include but are not limited to the following:

1. [50 + 25 increment/additional G] probe has 4-G stretch in the INVADER oligonucleotide, probe, or stacker.

2. [70] probe has 5-base stretch containing position 1

3. [60] probe has 5-base stretch containing position 2

4. [90] two or three base sequence repeats at least four times starting at position +1 in the probe

5. [100] probe hybridizing region is short (8 bases with a stacker or ≤ 12 bases without a stacker)

6. [40 + 5 increment/base] probe hybridizing region is long (≥ 17 bases with a stacker or ≥ 20 bases without a stacker)

7. [100] penultimate 3′ base of the INVADER oligonucleotide matches the 3′ base of the probe arm

In some embodiments, penalties are assessed for location of SNP variations at or near the cleavage site. In other embodiments, penalties are assessed based on cleavage site base preferences (*e.g.*, some enzyme may cleave after more efficiently after particular bases, such as Gs, and penalties may be used when a different base is placed in that location). In still other embodiments, penalties are assessed based on ranking of stacking interactions between a probe 3′ base and a stacking oligonucleotide 5′ base (*e.g.*, in some embodiments, AA stacks may perform better than TT stacks.

In particularly preferred embodiments, temperatures for each of the oligonucleotides in the designs are recomputed and scores are recomputed as changes are made. In some embodiments, score descriptions can be seen by clicking a "descriptions" button. In some embodiments, a BLAST search option is provided. In preferred embodiments, a BLAST search is done by clicking a "BLAST Design" button. In some embodiments, this action brings up a dialog box describing the BLAST process. In preferred embodiments, the BLAST search results are displayed as a highlighted design on a Designer Worksheet.

In some embodiments, a user accepts a design by clicking an "Accept" button. In other embodiments, the program approves a design without user intervention. In preferred embodiments, the program sends the approved design to a next process step (*e.g.*, into production; into a file or database). In some embodiments, the program provides a screen view (*e.g.*, an Output Page, Figure 10 OLD NUMBER), allowing review of the final designs created and allowing notes to be attached to the design. In preferred embodiments, the user can return to the Designer Worksheet (*e.g.*, by clicking a "Go Back" button) or can save the design (*e.g.*, by clicking a "Save It" button) and continue (*e.g.*, to submit the designed oligonucleotides for production).

In some embodiments, the program provides an option to create a screen view of a design optimized for printing (*e.g.*, a text-only view) or other export (*e.g.*, an Output view, Figure 11). In preferred embodiments, the Output view provides a description of the design particularly suitable for printing, or for exporting into another application (*e.g.*, by copying and pasting into another application). In particularly preferred embodiments, the Output view opens in a separate window.

One embodiments of a design session using the RIC module for RNA assay design is represented in Figure 13. The RIC module is shown by way of example; similar steps are followed in the SIC and TIC design modules represented in Figures 12 and 14, respectively. RNA assay design in this embodiment of the RIC module may comprise the following steps::

- entry of assay information into defined fields (e.g., user, assay name, assay abbreviation, etc.) (Figure 13A).
- user selects species via drop down menu (Figure 13B).
- user selects the RNA design module via RIC button (Figure 13A).
- RNA sequences (including FASTA format) is copied and pasted in (Figure 13C).
- cleavage site based design is indicated (*e.g.*, sites indicated are splice junctions, SNPs, or other any other sites selected by user, for example, using the bioinformatics assessment described above; user can enter multiple sites) (Figure 13C). Multiple probes can be designed per cleavage site (e.g., 257[3] gives three probes for the design for the 257 site).
- Stacking oligonucleotide design format can be selected (*e.g.*, "Has Stacker" button, Figure 13C).
- The user can change the non-complementary 5′ arm on the probe via a drop-down menu (Figure 13D).
- Bases can be added to or deleted from the 5′ end of the INVADER oligonucleotide(Figure 13E), the 3′ end of the probe (automatically adjusts stacking oligonucleotide position and length to satisfy it temperature setting) (Figure 13F), and the 3′ end of the stacking oligonucleotide.
- On the active design page the user can alter the INVADER oligonucleotide, probe, and stacking oligonucleotide temperatures (*e.g.*, Fig 13G). Exemplary default settings and actual calculated values are shown (*e.g.*, in a separate window).
- On the active design page the user can alter the target, INVADER oligonucleotide, probe, and stacking oligonucleotide concentrations *e.g.*, from default settings(Figure 13H);
- user can select enzymes (*e.g.*, alternative CLEAVASE enzymes) via drop-down menu.
- All input cleavage site designs can be shown on the same active design page (Figures 13D-H);

- and the user can select "Cancel" to go back to a previous screen. When finished making any adjustments to the designs, the user can select the "Design Review" button to get to the Design Review step. Design Review shows all entered assay information, the complete mRNA sequence (5′ to 3′), and the designed INVADER oligonucleotide set for each cleavage site aligned to its corresponding mRNA sequence (displayed here 3′ to 5′) (Figure 13I);
- synthetic target sequences are automatically generated including T7 promoter sequence that would enable generation of the mini-*in vitro* RNA transcript via a transcription kit and a mixture of the two synthetic target sequences.(*e.g.*, Figure 13I). Arrestor oligonucleotides are automatically designed for each probe and are fully complementary to the target-specific region of the probe and extend 6 nucleotides into the non-complementary 5′ arm. They appear in the INVADERCREATOR output file and are automatically ordered with all 2′-Ome bases (*e.g.*, Figure 13I);
- an "All" button can be selected to automatically order all oligonucleotides for a given design or individual oligonucleotides can be selected or deselected as desired, and a "Notes" field allows the user to type in any comments related to that particular design.
- The user selects either the "Job Submit" or "Printable Page/Job Submit" button to move on to the oligo ordering screen (Figure 13I).
- The user gets a listing of all oligonucleotides that were checked for ordering in the Design Review screen and selects each one to call up the oligo order form for that particular oligonucleotide (Figure 13J).
- An Oligo Request form is queued up for each oligo and the user has the ability to select an oligo type via a drop-down menu, the synthesis scale, purification method, various 5′, 3′, or internal modifications, the ability to select "Other" and input unique modifications not listed in the drop-down menus, the ability to highlight a portion of the sequence and designate and alternative nucleotide chemistry (e.g., 2'-Ome's or phosphorothioates) (13 L-0). In some embodiments, the software is set to automatically accept default values and submit all orders directly from the Design Review screen (*e.g.*, via n "order Oligonucleotides Now" button) without user review of an Oligo Request form.

- The user selects the "Submit to Synthesis" button when finished modifying a particular Oligo Request form and then queues up the remaining oligonucleotides in the order one by one and does likewise.

In some embodiments, the RIC module also allows the selection of multiple designs for one cleavage site. For example, entering "257, 257, 257, 512" in the sites box (*e.g.*, on Figure 13C for 13P) would give the same three designs for 257 and one for 512. As shown in 13P, one could also enter 257 [2] to create 2 designs to the 257 site. In some embodiments, the user has the ability to modify each design individually in the following steps.

One embodiment of a design session using the TIC module for RNA assay design is represented in Figure 14.

- This is the very first screen of automated order entry, and is the same regardless the format (SNP, RNA, Transgene. To go to Transgene InvaderCreator, click on the "TIC button (Figure 14A).
- In this screen the user can paste the Transgene or Internal control sequence. By filling out a number in the "number of loci" field, the user can choose how many designs he or she wants to see. The number of loci are evenly divided over the entered sequence. In addition to these loci, other cleavage sites can be indicated by bracketing a certain base "[C]". Also, by inserting a number before the base in the bracketed base, multiple probe arm designs can be made (*e.g.* "[3C]" would design 3 probes for site "C", each of which can have its own arm (Figure 14B).
- In this screen all the cleavage sites are shown (in sense and antisense orientation). The score is based on penalty scores also used in SNP IC. A perfect design has score 100. When both sense and antisense have a score of 100, a tiebreaker rule gives the winner one extra point. The computer program automatically picks the top two designs based on score, however the user can override those choices. (Figure 14C).

- This is the design page. In principal it is the same as the SNP Invader creator, with the exception that instead of having a sense and antisense design, you have a 1st and 2nd choice design. (Figure 14D).
- Once the designs have been optimized (ie. bases added or deleted) the user can go to the design review page. From here the oligos can be checked for automatic ordering.This is the top half of that page, the bottom half is on the next slide (14E-F).

C. RNA INVADER assay design.

For each design method, typically three different INVADER oligonucleotide sets would be designed and screened and the best performing set would be selected as the product assay. If sufficient detection was not achieved with the initial 3-site screen, a redesign method could include moving the cleavage site/accessible site 1 or more nucleotides in either direction and/or lower scoring designs not ordered in the initial process could be ordered and tested.

Integration of the various design methods could involve querying the user or having the user select one or more design methods based on the following examples:

Does the mRNA sequence have significant homology to other genes or gene family members? If yes, should the target sequence be detected exclusively or inclusively?

Is the mRNA sequence one of 2 or more alternatively spliced variants? If yes, should the target sequence be detected exclusively or inclusively?

If closely related sequences or alternatively spliced variants are not identified in the sequence analysis (e.g., via the bioinformatics module), should the candidate assays be designed via the splice site or accessible site method?

Alternatively, as described above, these types of questions can be encoded in an algorithm that would automatically determine the best design strategy based on the automated sequence analysis in the bioinformatics module.

Splice site design. If assay specificity and/or performance requirements do not dictate otherwise, assays can be designed at or near splice junctions to completely preclude the possibility of detecting genomic DNA in a sample. Splice site design involves determining the splice junctions within the mRNA, usually via pairwise alignment of the mRNA sequence with the genomic DNA sequence for that gene, and then locating INVADER assay cleavage sites at or near the splice site. Typically, the INVADER oligonucleotide is positioned on one side of the splice junction and the probe and stacking oligonucleotide (if used) are positioned on the other side. Thus, if the oligonucleotides were bound to genomic DNA, the probe and INVADER oligonucleotides would be separated by the intervening intronic sequences, which would preclude formation of the required overlap substrate for the CLEAVASE enzyme.

Accessible site design. Again, if assay specificity and/or performance requirements do not dictate otherwise, assays can also be designed to accessible sites within the mRNA. Accessible sites are unstructured regions of the RNA and those determined experimentally, for example, using RT-ROL (Allawi *et al.* RNA 7:314 [2001]), usually correlate well with enhanced INVADER RNA assay performance. Accessible sites can also be determined via *in silico* analysis. For example, the RNA sequence could be folded in m-Fold software and then analyzed in Oligowalk to determine accessible sites in the RNA. A program could be written to automatically output the accessible sites (defined as a region with negative Overall G values for an oligonucleotide binding to that region) for the folded RNA. For example, the program could determine when there were 5 or more consecutive nucleotides with Overall G values of –5 or less, then determine the midpoint of this region, and then output those sites into a file. For example, a 10-base negative G region encompassing target sequence nucleotides 200-210 would correspond to an accessible site at 205.

In either case, accessible site design could be encoded into the INVADERCREATOR module by method A or B.

Method A

Assays could be designed in reverse of the cleavage site design process. The user would specify the precise position of the 3' end of the probe within an accessible site and the probe would be built out toward the 5' end to satisfy the preset Tm requirement. Stacking

oligonucleotide (if designing in a stacker format) contributions to the probe's Tm would be determined as the probe was being built and the Invader oligonucleotide would be designed after the program finished the probe or probe/stacker design.

Method B

Hatim Allawi suggested an alternative method for accessible site design that could use the same probe-building algorithm that is used for cleavage site design methods. The user could enter the accessible site and the INVADERCREATOR module could shift a defined number of bases (a default shift could be determined) downstream. For example, 200 could be entered as an accessible site, and INVADERCREATOR module would build a design using the existing algorithm for cleavage site 210 if the shift value was 10. Next to the check box for "Stacker Design" could be a check box for "Accessible Site Design". Next to this check box could be a field in which the user would designate the number of bases to shift. The current "Cleavage Sites" field could say "Design Sites" to generically encompass either design mode (cleavage sites or accessible sites). Users could have the capability to check one or both boxes (e.g. stacker design and accessible site design, accessible site design only, etc.).

Splice variant design. Splice variant assays can be designed in a variety of ways. An inclusive detection assay could be designed to detect a region of sequence (e.g. a particular exon) present in all variants. A particular splice variant could be detected by designing the assay to a unique splice site (e.g. if a 5 exon gene yields a splice variant that excludes exon 3, the assay could be designed to detect the exon 2-exon 4 splice junction). Since specificity of the INVADER RNA assay is primarily linked to discrimination at the cleavage site, even very small exonic sequences (e.g. a few nucleotides) could be distinguished. In some cases, it may be useful to detect not any one particular mRNA variant but to individually quantitate exons and/or splice junctions in a pool of mRNA variants. The quantitation pattern from this type of INVADER RNA assay analysis may correlate with particular cellular processes or metabolic states.

Discrimination site design. Closely-related sequences would be aligned to the input target sequence and an automated analysis could be performed to identify all sites that contain, for example, two or more adjacent base differences for any one sequence from all others in the alignment. Another automated analysis algorithm could determine regions of homology of sufficient size to accommodate an INVADER oligonucleotide probe set that would inclusively detect all closely-related mRNAs. An output of the location of such double base discrimination sites or regions of homology could be reviewed by the user before accessing the INVADERCREATOR module or automatically designed via input of a batch file.

The present invention is not limited to the use of the INVADERCREATOR software. Indeed, a variety of software programs are contemplated and are commercially available, including, but not limited to GCG Wisconsin Package (Genetics computer Group, Madison, WI) and Vector NTI (Informax, Rockville, Maryland).

In some embodiments, the present invention provides design parameters for combining multiple nucleic acid detection technologies. For example, in some embodiments, INVADER assays or other assays are used in conjunction with amplified nucleic acid obtained by using the polymerase chain reaction (PCR). In some preferred embodiments, PCR is run simultaneously with other assays.

D. TAQMAN Probe and Primer Design

A number of different strategies can be used to design TaqMan (5' Nuclease assay) Probes. The following are example of considerations that may be used when designing TAQMAN probes. One consideration is to design PCR primers such that the amplicon size is between 50 – 150 base pairs. Another consideration is to design PCR primers that have a Tm of around 60°C, with less than 2°C difference in Tm between forward and reverse primers. Preferred primers have GC% around 40-60% and have three or less consecutive runs of any nucleotide. Preferably, the primers have total lengths of between 18-25 nucleotides in length. PCR Primers are designed to have minimal haripin and minimal dimer formation tendencies (See below). Following selection of the PCR primers, the TAQMAN probe is then chosen from within the amplicon region, and has a Tm of about 10°C higher than the Tm of the PCR primers (typically, 70°C). TAQMAN probes should have a 5' FAM and a 3' TAMRA (or other labels),

and not begin with G. TAQMAN probes may be chosen, for example, by using programs such as OligoWalk to scan through the amplicon sequence and a probe chosen based upon predicted most stable thermodynamic parameters. Moreover, candidate TAQMAN probes can be eliminated which forms more than three consecutive basepairs with the PCR primers.

E. Multiplex PCR Primer Design

The INVADER assay can be used for the detection of single nucleotide polymorphisms (SNPs) with as little as 100-10 ng of genomic DNA without the need for target pre-amplification. However, with more than 80,000 INVADER assays developed and the potential for whole genome association studies involving hundreds of thousands of SNPs, the amount of sample DNA becomes a limiting factor for large-scale analysis. Due to the sensitivity of the INVADER assay on human genomic DNA (hgDNA) without target amplification, multiplex PCR coupled with the INVADER assay requires only limited target amplification (10^3-10^4) as compared to typical multiplex PCR reactions that require extensive amplification (10^9-10^{12}) for conventional gel detection methods. The low level of target amplification used for INVADER assay detection provides for more extensive multiplexing by avoiding amplification inhibition commonly resulting from target accumulation.

In some embodiments, it may be desired to detect related loci in a multiplex PCR reaction. In some such embodiments, the similarity between loci may prevent or complicate detection assay analysis of the sequence, as the detection assay technology may not be able to sufficiently discriminate between the closely related sequences. The present invention provides methods to overcome such problems, by generating a unique target sequence using a nucleic acid amplification technique (e.g., PCR), such that the unique target sequence is tested by the detection assay, rather the original sample (e.g., genomic DNA). This method is compatible with multiplexing, where considerations are made to ensure that amplified target sequence meets several criteria: 1) that the target sequence contains the polymorphism to be analyzed; 2) that the target sequence represents a unique target sequence (i.e., it is the only sequence in the reaction mixture that is detected by a detection assay designed to target the target sequence); and 3) that the target sequence does not contain other polymorphisms that are detected by any of the detection assays present in the multiplex reaction. Suitable detection assay components may be

selected with methods similar to those described above for the INVADERCREATOR methods. For example, in some embodiments, the software performs a BLAST alignment of the target sequence used for the SNP assay to find similar sequences in the genome that may generate the cross-reactivity signal. The design of PCR primers with software program should prevent amplification of any of the similar loci except the locus containing the SNP. To avoid pre-amplification of sequences other than the specific SNP sequence, the software performs a BLAST alignment of the sequence amplified with a pair of primers against all other detection assay sequences included in the pool. If cross-reactivity or potential cross-reactivity exists, the set of primers is redesigned or the co-amplified sequences are included in different pools.

The same type of design analysis may be used for detection assays directed at the detection of haplotypes. For example, primers are generated to amplify sets of target sequences that each uniquely contain the polymorphisms to be detected.

In some embodiments, multiplex detection assays are provided in a plurality of arrays. For example, in some embodiments, a first array comprises assays configured for detection directly from genomic DNA and a second array comprises assays configured for pre-amplification of target sequences from genomic DNA prior to detection assay analysis of the target sequence.

In some preferred embodiments, only limited pre-amplification of target sequences is carried out prior to detection by the detection assay. For example, in some embodiments, only a 10^5-10^6 fold or less increase in target copy number is obtained prior to detection. This is in contrast to typical PCR reactions where 10^{10}-10^{12} or more fold amplification is utilized in detection reactions. In certain embodiments, 100 genotypes from a single PCR amplification are possible with the methods and systems of the present invention using only 10 ng of genomic DNA (e.g. less than 0.1 ng of human genomic DNA per SNP).

In some embodiments, kits are provided for pre-amplification and detection of target sequences. In some embodiments, the kits comprise amplification primers. For multiplex reactions, the amplification primers may be provided in a single container. The amplification primers may also be packaged with detection assay components. In some embodiments, amplification primers and detection assay components (e.g., INVADER assay components) are provided in a single container (e.g., in a single well of a multiwell plate). In some embodiments,

the reaction components are provided in dry form in a reaction chamber. In some such embodiments, the kits are configured to allow reactions to occur where the only thing that is added to the reaction chamber is a solution containing genomic DNA.

The present invention provides methods and selection criteria that allow primer sets for multiplex PCR to be generated (e.g. that can be coupled with a detection assay, such as the INVADER assay). In some embodiments, software applications of the present invention automated multiplex PCR primer selection, thus allowing highly multiplexed PCR with the primers designed thereby. Using the INVADER Medically Associated Panel (MAP) as a corresponding platform for SNP detection, as shown in PCR primer example 2 (below), the methods, software, and selection criteria of the present invention allowed accurate genotyping of 94 of the 101 possible amplicons (~93%) from a single PCR reaction. The original PCR reaction used only 10 ng of hgDNA as template, corresponding to less than 150 pg hgDNA per INVADER assay.

The multiplex primer design systems may be employed to design PCR primer sets useful with a particular type of assay, such as the INVADER assay. Figure 15 illustrates creation of one of the primer pairs (both a forward and reverse primer) for a 101 primer set from sequences available for analysis on the INVADER Medically Associated Panel using one embodiment of the software application of the present invention. Figure 15A shows a sample input file of a single entry (e.g. shows target sequence information for a single target sequence containing a SNP that is processed the method and software of the present invention). The target sequence information in Figure 15 includes Third Wave Technologies's SNP#, short name identifier, and sequence with the SNP location indicated in brackets. Figure 15B shows the sample output file of a the same entry (e.g. shows the target sequence after being processed by the systems and methods and software of the present invention. The output information includes the sequence of the footprint region (capital letters flanking SNP site, showing region where INVADER assay probes hybridize to this target sequence in order to detect the SNP in the target sequence), forward and reverse primer sequences (bold), and their corresponding Tm's.

In some embodiments, the selection of primers to make a primer set capable of multiplex PCR is performed in automated fashion (e.g. by a software application). Automated primer

selection for multiplex PCR may be accomplished employing a software program designed as shown by the flow chart in Figure 17.

Multiplex PCR commonly requires extensive optimization to avoid biased amplification of select amplicons and the amplification of spurious products resulting from the formation of primer-dimers. In order to avoid these problems, the present invention provides methods and software application that provide selection criteria to generate a primer set configured for multiplex PCR, and subsequent use in a detection assay (e.g. INVADER detection assays).

In some embodiments, the methods and software applications of the present invention start with user defined sequences and corresponding SNP locations. In certain embodiments, the methods and/or software application determines a footprint region within the target sequence (the minimal amplicon required for INVADER detection) for each sequence (shown in capital letters in Figure 15B). The footprint region includes the region where assay probes hybridize, as well as any user defined additional bases extending outward therefore (e.g. 5 additional bases included on each side of where the assay probes hybridize). Next, primers are designed outward from the footprint region and evaluated against several criteria, including the potential for primer-dimer formation with previously designed primers in the current multiplexing set (See, primers in bold in Figure 15A, and selection steps in Figure 17). This process may be continued, as shown in Figure 17, through multiple iterations of the same set of sequences until primers against all sequences in the current multiplexing set can be designed.

Once a primer set is designed for multiplex PCR, this set may be employed, in some embodiments, as shown in the basic workflow scheme shown in Figure 16. Multiplex PCR may be carried out, for example, under standard conditions using only 10 ng of hgDNA as template. After 10 min at 95°C, Taq (2.5 units) may be added to a 50ul reaction and PCR carried out for 50 cycles. The PCR reaction may be diluted and loaded directly onto an INVADER MAP plate (3ul/well) (See Figure 16). An additional 3ul of 15mM $MgCl_2$ may be added to each reaction on the INVADER MAP plate and covered with 6ul of mineral oil. The entire plate may then be heated to 95°C for 5 min. and incubated at 63°C for 40 min. FAM and RED fluorescence may then be measured on a Cytofluor 4000 fluorescent plate reader and "Fold Over Zero" (FOZ) values calculated for each amplicon. Results from each SNP may be color coded in a table as

"pass" (green), "mis-call" (pink), or "no-call" (white) (See, PCR Primer Design Example 2 below).

In some embodiments the number of PCR reactions is from about 1 to about 10 reactions. In some embodiments, the number of PCR reactions is from about 10 to about 50 reactions. In further embodiments, the number of PCR reactions is from about 50 to about 100. In additional embodiments, the number of PCR reactions is greater than 100.

The present invention also provides methods to optimize multiplex PCR reactions (e.g. once a primer set is generated, the concentration of each primer or primer pair may be optimized). For example, once a primer set has been generated and used in a multiplex PCR at equal molar concentrations, the primers may be evaluated separately such that the optimum primer concentration is determined such that the multiplex primer set performs better.

Multiplex PCR reactions are being recognized in the scientific, research, clinical and biotechnology industries as potentially time effective and less expensive means of obtaining nucleic acid information compared to standard, monoplex PCR reactions. Instead of performing only a single amplification reaction per reaction vessel (tube or well of a multi-well plate for example), numerous amplification reactions are performed in a single reaction vessel.

The cost per target is theoretically lowered by eliminating technician time in assay set-up and data analysis, and by the substantial reagent savings (especially enzyme cost). Another benefit of the multiplex approach is that far less target sample is required. In whole genome association studies involving hundreds of thousands of single nucleotide polymorphisms (SNPs), the amount of target or test sample is limiting for large scale analysis, so the concept of performing a single reaction, using one sample aliquot to obtain, for example, 100 results, versus using 100 sample aliquots to obtain the same data set is an attractive option.

To design primers for a successful multiplex PCR reaction, the issue of aberrant interaction among primers should be addressed. The formation of primer dimers, even if only a few bases in length, may inhibit both primers from correctly hybridizing to the target sequence. Further, if the dimers form at or near the 3' ends of the primers, no amplification or very low levels of amplification will occur, since the 3' end is required for the priming event. Clearly, the more primers utilized per multiplex reaction, the more aberrant primer interactions are possible.

The methods, systems and applications of the present help prevent primer dimers in large sets of primers, making the set suitable for highly multiplexed PCR.

When designing primer pairs for numerous sites (for example 100 sites in a multiplex PCR reaction), the order in which primer pairs are designed can influence the total number of compatible primer pairs for a reaction. For example, if a first set of primers is designed for a first target region that happens to be an A/T rich target region, these primers will be A/T rich. If the second target region chosen also happens to be an A/T rich target region, it is far more likely that the primers designed for these two sets will be incompatible due to aberrant interactions, such as primer dimers. If, however, the second target region chosen is not A/T rich, it is much more likely that a primer set can be designed that will not interact with the first A/T rich set. For any given set of input target sequences, the present invention randomizes the order in which primer sets are designed (See, Figure 17). Furthermore, in some embodiments, the present invention re-orders the set of input target sequences in a plurality of different, random orders to maximize the number of compatible primer sets for any given multiplex reaction (See, Figure 17). In certain embodiments, the primers are designed such that GC-rich and AT-rich regions are avoided.

The present invention provides criteria for primer design that minimizes 3' interactions (e.g. 3' complementarity of primers is avoided to reduce probability of primer-dimer formation), while maximizing the number of compatible primer pairs for a given set of reaction targets in a multiplex design. For primers described as 5'-N[x]-N[x-1]-.....-N[4]-N[3]-N[2]-N[1]-3', N[1] is an A or C (in alternative embodiments, N[1] is a G or T). N[2]-N[1] of each of the forward and reverse primers designed should not be complementary to N[2]-N[1] of any other oligonucleotide. In certain embodiments, N[3]-N[2]-N[1] should not be complementary to N[3]-N[2]-N[1] of any other oligonucleotide. In preferred embodiments, if these criteria are not met at a given N[1], the next base in the 5' direction for the forward primer or the next base in the 3' direction for the reverse primer may be evaluated as an N[1] site. This process is repeated, in conjunction with the target randomization, until all criteria are met for all, or a large majority of, the targets sequences (e.g. 95% of target sequences can have primer pairs made for the primer set that fulfill these criteria).

Another challenge to be overcome in a multiplex primer design is the balance between actual, required nucleotide sequence, sequence length, and the oligonucleotide melting temperature (Tm) constraints. Importantly, since the primers in a multiplex primer set in a reaction should function under the same reaction conditions of buffer, salts and temperature, they need therefore to have substantially similar Tm's, regardless of GC or AT richness of the region of interest. The present invention allows for primer design that meets minimum Tm and maximum Tm requirements and minimum and maximum length requirements. For example, in the formula for each primer 5'-N[x]-N[x-1]-.....-N[4]-N[3]-N[2]-N[1]-3', x is selected such the primer has a predetermined melting temperature (e.g. bases are included in the primer until the primer has a calculated melting temperature of about 50 degrees Celsius). In certain embodiments, each of the primers in a set has the same melting temperature.

Often the products of a PCR reaction are used as the target material for another nucleic acid detection means, such as a hybridization-type detection assays, or the INVADER reaction assays for example. Consideration should be given to the location of primer placement to allow for the secondary reaction to successfully occur, and again, aberrant interactions between amplification primers and secondary reaction oligonucleotides should be minimized for accurate results and data. Selection criteria may be employed such that the primers designed for a multiplex primer set do not react (e.g. hybridize with, or trigger reactions) with oligonucleotide components of a detection assay. For example, in order to prevent primers from reacting with the FRET oligonucleotide of a bi-plex INVADER assay, certain homology criteria is employed. In particular, if each of the primers in the set are defined as 5'-N[x]-N[x-1]-.....-N[4]-N[3]-N[2]-N[1]-3', then N[4]-N[3]-N[2]-N[1]-3' is selected such that it is less than 90% homologous with the FRET or INVADER oligonucleotides. In other embodiments, N[4]-N[3]-N[2]-N[1]-3' is selected for each primer such that it is less than 80% homologous with the FRET or INVADER oligonucleotides. In certain embodiments, N[4]-N[3]-N[2]-N[1]-3' is selected for each primer such that it is less than 70% homologous with the FRET or INVADER oligonucleotides.

While employing the criteria of the present invention to develop a primer set, some primer pairs may not meet all of the stated criteria (these may be rejected as errors). For example, in a set of 100 targets, 30 are designed and meet all listed criteria, however, set 31 fails. In the method of the present invention, set 31 may be flagged as failing, and the method could

continue through the list of 100 targets, again flagging those sets which do not meet the criteria (See Figure 17). Once all 100 targets have had a chance at primer design, the method would note the number of failed sets, re-order the 100 targets in a new random order and repeat the design process (See, Figure 17). After a configurable number of runs, the set with the most passed primer pairs (the least number of failed sets) are chosen for the multiplex PCR reaction (See Figure 17).

Figure 17 shows a flow chart with the basic flow of certain embodiments of the methods and software application of the present invention. In preferred embodiments, the processes detailed in Figure 17 are incorporated into a software application for ease of use (although, the methods may also be performed manually using, for example, Figure 17 as a guide).

Target sequences and/or primer pairs are entered into the system shown in Figure 17. The first set of boxes show how target sequences are added to the list of sequences that have a footprint determined (See "B" in Figure 17), while other sequences are passed immediately into the primer set pool (e.g. PDPass, those sequences that have been previously processed and shown to work together without forming Primer dimers or having reactivity to FRET sequences), as well as DimerTest entries (e.g. pair or primers a user wants to use, but that has not been tested yet for primer dimer or fret reactivity). In other words, the initial set of boxes leading up to "end of input" sort the sequences so they can be later processed properly.

Starting at "A" in Figure 17, the primer pool is basically cleared or "emptied" to start a fresh run. The target sequences are then sent to "B" to be processed, and DimerTest pairs are sent to "C" to be processed. Target sequences are sent to "B", where a user or software application determines the footprint region for the target sequence (e.g. where the assay probes will hybridize in order to detect the mutation (e.g. SNP) in the target sequence). This region is generally shown in capital letters in figures, such as Figure 15B. It is important to design this region (which the user may further expand by defining that additional bases past the hybridization region be added) such that the primers that are designed fully encompass this region. In Figure 17, the software application INVADER CREATOR is used to design the INVADER oligonucleotide and downstream probes that will hybridize with the target region (although any type of program of system could be used to create any type of probes a user was interested in designing probes for, and thus determining the footprint region for on the target

sequence). Thus the core footprint region is then defined by the location of these two assay probes on the target.

Next, the system starts from the 5' edge of the footprint and travels in the 5' direction until the first base is reached, or until the first A or C (or G or T) is reached. This is set as the initial starting point for defining the sequence of the forward primer (i.e. this serves as the initial N[1] site). From this initial N[1] site, the sequence of the primer for the forward primer is the same as those bases encountered on the target region. For example, if the default size of the primer is set as 12 bases, the system starts with the bases selected as N[1] and then adds the next 11 bases found in the target sequences. This 12-mer primer is then tested for a melting temperature (e.g. using INVADER CREATOR), and additional bases are added from the target sequence until the sequence has a melting temperature that is designated by the user (e.g. about 50 degrees Celsius, and not more than 55 degrees Celsius). For example, the system employs the formula 5'-N[x]-N[x-1]-.....-N[4]-N[3]-N[2]-N[1]-3', and x is initially 12. Then the system adjusts x to a higher number (e.g. longer sequences) until the pre-set melting temperature is found.

The next box in Figure 17, is used to determine if the primer that has been designed so far will cause primer-dimer and/or fret reactivity (e.g. with the other sequences already in the pool). The criteria used for this determination are explained above. If the primer passes this step, the forward primer is added to the primer pool. However, if the forward primer fails this criteria, as shown in Figure 17, the starting point (N[1] is moved) one nucleotide in the 5' direction (or to the next A or C, or next G or T). The system first checks to make sure shifting over leaves enough room on the target sequence to successfully make a primer. If yes, the system loops back and check this new primer for melting temperature. However, if no sequence can be designed, then the target sequence is flagged as an error (e.g. indicating that no forward primer can be made for this target).

This same process is then repeated for designing the reverse primer, as shown in Figure 17. If a reverse primer is successfully made, then the pair or primers is put into the primer pool, and the system goes back to "B" (if there are more target sequences to process), or goes onto "C" to test DimerTest pairs.

Starting a "C" in Figure 17 shows how primer pairs that are entered as primers (DimerTest) are processed by the system. If there are no DimerTest pairs, as shown in Figure 17, the system goes on to "D". However, if there are DimerTest pairs, these are tested for primer-dimer and/or FRET reactivity as described above. If the DimerTest pair fails these criteria they are flagged as errors. If the DimerTest pair passes the criteria, they are added to the primer set pool, and then the system goes back to "C" if there are more DimerTest pairs to be evaluated, or goes on to "D" if there are no more DimerTest pairs to be evaluated.

Starting at "D" in Figure 17, the pool of primers that has been created is evaluated. The first step in this section is to examine the number of error (failures) generated by this particular randomized run of sequences. If there were no errors, this set is the best set as maybe outputted to a user. If there are more than zero errors, the system compares this run to any other previous runs to see what run resulted in the fewest errors. If the current run has fewer errors, it is designated as the current best set. At this point, the system may go back to "A" to start the run over with another randomized set of the same sequences, or the pre-set maximum number of runs (e.g. 5 runs) may have been reached on this run (e.g. this was the 5th run, and the maximum number of runs was set as 5). If the maximum has been reached, then the best set is outputted as the best set. This best set of primers may then be used to generate as physical set of oligonucleotides such that a multiplex PCR reaction may be carried out.

Another challenge to be overcome with multiplex PCR reactions is the unequal amplicon concentrations that result in a standard multiplex reaction. The different loci targeted for amplification may each behave differently in the amplification reaction, yielding vastly different concentrations of each of the different amplicon products. The present invention provides methods, systems, software applications, computer systems, and a computer data storage medium that may be used to adjust primer concentrations relative to a first detection assay read (e.g. INVADER assay read) , and then with balanced primer concentrations come close to substantially equal concentrations of different amplicons. A generalized protocol for such multiplex optimization is presented in Figure 17.

The concentrations for various primer pairs may be determined experimentally. In some embodiments, there is a first run conducted with all of the primers in equimolar concentrations. Time reads are then conducted. Based upon the time reads, the relative amplification factors for

each amplicon are determined. Then based upon a unifying correction equation, an estimate of what the primer concentration should be obtained to get the signals closer within the same time point. These detection assays can be on an array of different sizes (384 well plates).

It is appreciated that combining the invention with detection assays and arrays of detection assays provides substantial processing efficiencies. Employing a balanced mix of primers or primer pairs created using the invention, a single point read can be carried out so that an average user can obtain great efficiencies in conducting tests that require high sensitivity and specificity across an array of different targets.

Having optimized primer pair concentrations in a single reaction vessel allows the user to conduct amplification for a plurality or multiplicity of amplification targets in a single reaction vessel and in a single step. The yield of the single step process is then used to successfully obtain test result data for, for example, several hundred assays. For example, each well on a 384 well plate can have a different detection assay thereon. The results of the single step mutliplex PCR reaction has amplified 384 different targets of genomic DNA, and provides you with 384 test results for each plate. Where each well has a plurality of assays even greater efficiencies can be obtained.

Therefore, the present invention provides the use of the concentration of each primer set in highly multiplexed PCR as a parameter to achieve an unbiased amplification of each PCR product. Any PCR includes primer annealing and primer extension steps. Under standard PCR conditions, high concentration of primers in the order of 1 uM ensures fast kinetics of primers annealing while the optimal time of the primer extension step depends on the size of the amplified product and can be much longer than the annealing step. By reducing primer concentration, the primer annealing kinetics can become a rate limiting step and PCR amplification factor should strongly depend on primer concentration, association rate constant of the primers, and the annealing time.

The binding of primer P with target T can be described by the following model:

$$P + T \xrightarrow{k_a} PT \quad (1)$$

where k_a is the association rate constant of primer annealing. We assume that the annealing occurs at the temperatures below primer melting and the reverse reaction can be ignored.

The solution for this kinetics under the conditions of a primer excess is well known:

$$[PT] = T_0(1 - e^{-k_a ct}) \quad (2)$$

where $[PT]$ is the concentration of target molecules associated with primer, T_0 is initial target concentration, c is the initial primer concentration, and t is primer annealing time. Assuming that each target molecule associated with primer is replicated to produce full size PCR product, the target amplification factor in a single PCR cycle is

$$Z = \frac{T_0 + [PT]}{T_0} = 2 - e^{-k_a ct} \quad (3)$$

The total PCR amplification factor after n cycles is given by

$$F = Z^n = (2 - e^{-k_a ct})^n \quad (4)$$

As it follows from equation 4, under the conditions where the primer annealing kinetics is the rate limiting step of PCR, the amplification factor should strongly depend on primer concentration. Thus, biased loci amplification, whether it is caused by individual association rate constants, primer extension steps or any other factors, can be corrected by adjusting primer concentration for each primer set in the multiplex PCR. The adjusted primer concentrations can be also used to correct biased performance of INVADER assay used for analysis of PCR pre-amplified loci. Employing this basic principle, the present invention has demonstrated a linear relationship between amplification efficiency and primer concentration and used this equation to balance primer concentrations of different amplicons, resulting in the equal amplification of ten different amplicons in PCR Primer Design Example 1. This technique may be employed on any size set of multiplex primer pairs. In some embodiments, the PCR primers are unoptimized, and the INVADER assay is employed to detect the amplified products (See, Ohnishi et al., J. Hum. Genet. 46:471-7, 2001, herein incorporated by reference.

i. PCR Primer Design Example 1

The following experimental example describes the manual design of amplification primers for a multiplex amplification reaction, and the subsequent detection of the amplicons by the INVADER assay.

Ten target sequences were selected from a set of pre-validated SNP-containing sequences, available in a TWT in-house oligonucleotide order entry database (see Figure 18). Each target contains a single nucleotide polymorphism (SNP) to which an INVADER assay had

been previously designed. The INVADER assay oligonucleotides were designed by the INVADER CREATOR software (Third Wave Technologies, Inc. Madison, WI), thus the footprint region in this example is defined as the INVADER "footprint", or the bases covered by the INVADER and the probe oligonucleotides, optimally positioned for the detection of the base of interest, in this case, a single nucleotide polymorphism (See Figure 18). About 200 nucleotides of each of the 10 target sequences were analyzed for the amplification primer design analysis, with the SNP base residing about in the center of the sequence. The sequences are shown in Figure 18.

Criteria of maximum and minimum probe length (defaults of 30 nucleotides and 12 nucleotides, respectively) were defined, as was a range for the probe melting temperature Tm of 50- 60°C. In this example, to select a probe sequence that will perform optimally at a pre-selected reaction temperature, the melting temperature (T_m) of the oligonucleotide is calculated using the nearest-neighbor model and published parameters for DNA duplex formation (Allawi and SantaLucia, Biochemistry, 36:10581 [1997], herein incorporated by reference). Because the assay's salt concentrations are often different than the solution conditions in which the nearest-neighbor parameters were obtained (1M NaCl and no divalent metals), and because the presence and concentration of the enzyme influence optimal reaction temperature, an adjustment should be made to the calculated T_m to determine the optimal temperature at which to perform a reaction. One way of compensating for these factors is to vary the value provided for the salt concentration within the melting temperature calculations. This adjustment is termed a 'salt correction'. The term "salt correction" refers to a variation made in the value provided for a salt concentration for the purpose of reflecting the effect on a T_m calculation for a nucleic acid duplex of a non-salt parameter or condition affecting said duplex. Variation of the values provided for the strand concentrations will also affect the outcome of these calculations. By using a value of 280nM NaCl (SantaLucia, Proc Natl Acad Sci U S A, 95:1460 [1998], herein incorporated by reference) and strand concentrations of about 10 pM of the probe and 1 fM target, the algorithm for used for calculating probe-target melting temperature has been adapted for use in predicting optimal primer design sequences.

Next, the sequence adjacent to the footprint region, both upstream and downstream were scanned and the first A or C was chosen for design start such that for primers described as 5'-

N[x]-N[x-1]-.....-N[4]-N[3]-N[2]-N[1]-3', where N[1] should be an A or C. Primer complementarity was avoided by using the rule that: N[2]-N[1] of a given oligonucleotide primer should not be complementary to N[2]-N[1] of any other oligonucleotide, and N[3]-N[2]-N[1] should not be complementary to N[3]-N[2]-N[1] of any other oligonucleotide. If these criteria were not met at a given N[1], the next base in the 5' direction for the forward primer or the next base in the 3' direction for the reverse primer will be evaluated as an N[1] site. In the case of manual analysis, A/C rich regions were targeted in order to minimize the complementarity of 3' ends.

In this example, an INVADER assay was performed following the multiplex amplification reaction. Therefore, a section of the secondary INVADER reaction oligonucleotide (the FRET oligonucleotide sequence) was also incorporated as criteria for primer design; the amplification primer sequence should be less than 80% homologous to the specified region of the FRET oligonucleotide.

The output primers for the 10-plex multiplex design are shown in Figure 18). All primers were synthesized according to standard oligonucleotide chemistry, desalted (by standard methods) and quantified by absorbance at A260 and diluted to 50 μM concentrated stock. Multiplex PCR was then carried out using 10-plex PCR using equimolar amounts of primer (0.01 uM/primer) under the following conditions; 100 mM KCl, 3 mM MgCl, 10mM Tris pH8.0, 200uM dNTPs, 2.5U taq, and 10ng of human genomic DNA (hgDNA) template in a 50ul reaction. The reaction was incubated for (94C/30sec, 50C/44sec.) for 30 cycles. After incubation, the multiplex PCR reaction was diluted 1:10 with water and subjected to INVADER analysis using INVADER Assay FRET Detection Plates, 96 well genomic biplex, 100ng CLEAVASE VIII, INVADER assays were assembled as 15ul reactions as follows; 1ul of the 1:10 dilution of the PCR reaction, 3ul of PPI mix, 5ul of 22.5 mM MgCl2, 6ul of dH20, covered with 15ul of Chillout. Samples were denatured in the INVADER biplex by incubation at 95C for 5min., followed by incubation at 63C and fluorescence measured on a Cytofluor 4000 at various timepoints.

Using the following criteria to accurately make genotyping calls (FOZ_FAM+FOZ_RED-2 > 0.6), only 2 of the 10 INVADER assay calls can be made after 10 minutes of incubation at 63C, and only 5 of the 10 calls could be made following an additional

50 min of incubation at 63C (60 min.) (See, Figure 19A). At the 60 min time point, the variation between the detectable FOZ values is over 100 fold between the strongest signal (Figure 19A, 41646, FAM_FOZ+RED_FOZ-2=54.2, which is also is far outside of the dynamic range of the reader) and the weakest signal (Figure 19A, 67356, FAM_FOZ+RED_FOZ-2=0.2). Using the same INVADER assays directly against 100ng of human genomic DNA (where equimolar amounts of each target would be available), all reads could be made with in the dynamic range of the reader and variation in the FOZ values was approximately seven fold between the strongest (Figure 19, 53530, FAM_FOZ+RED_FOZ-2=3.1) and weakest (Figure 19, 53530, FAM_FOZ+RED_FOZ-2=0.43) of the assays. This suggests that the dramatic discrepancies in FOZ values seen between different amplicons in the same multiplex PCR reaction is a function of biased amplification, and not variability attributable to INVADER assay. Under these conditions, FOZ values generated by different INVADER assays are directly comparable to one another and can reliably be used as indicators of the efficiency of amplification.

Estimation of amplification factor of a given amplicon using FOZ values. In order to estimate the amplification factor (*F*) of a given amplicon, the FOZ values of the INVADER assay can be used to estimate amplicon abundance. The FOZ of a given amplicon with unknown concentration at a given time (FOZm) can be directly compared to the FOZ of a known amount of target (e.g. 100 ng of genomic DNA = 30,000 copies of a single gene) at a defined point in time (FOZ_{240}, 240 min) and used to calculate the number of copies of the unknown amplicon. In equation 1, *FOZm* represents the sum of RED_FOZ and FAM_FOZ of an unknown concentration of target incubated in an INVADER assay for a given amount of time (*m*). FOZ_{240} represents an empirically determined value of RED_FOZ (using INVADER assay 41646), using for a known number of copies of target (e.g. 100ng of hgDNA ≅ 30,000 copies) at 240 minutes.

$$F = ((FOZ_m - 1) * 500 / (FOZ_{240} - 1)) * (240 / m)^{\wedge}2 \text{ (equation 1a)}$$

Although equation 1a is used to determine the linear relationship between primer concentration and amplification factor *F*, equation 1a' is used in the calculation of the amplification factor *F* for the 10-plex PCR (both with equimolar amounts of primer and optimized concentrations of primer), with the value of *D* representing the dilution factor of the PCR reaction. In the case of a 1:3 dilution of the 50 ul multiplex PCR reaction. *D*=0.3333.

$$F = ((FOZ_m - 2) * 500 / (FOZ_{240} - 1) * D) * (240 / m)^{\wedge}2 \text{ (equation 1a')}$$

Although equations 1a and 1a' will be used in the description of the 10-plex multiplex PCR, a more correct adaptation of this equation was used in the optimization of primer concentrations in the 107 plex PCR. In this case, FOZ_{240}=the average of FAM_FOZ_{240}+RED_FOZ_{240} over the entire INVADER MAP plate using hgDNA as target (FOZ_{240}=3.42) and the dilution factor D is set to 0.125.

$$F = ((FOZ_m - 2) * 500 / (FOZ_{240} - 2) * D) * (240 / m)\text{^}2 \text{ (equation 1b)}$$

It should be noted that in order for the estimation of amplification factor F to be more accurate, FOZ values should be within the dynamic range of the instrument on which the reading are taken. In the case of the Cytofluor 4000 used in this study, the dynamic range was between about 1.5 and about 12 FOZ.

Section 3. Linear Relationship between Amplification Factor and Primer Concentration.

In order to determine the relationship between primer concentration and amplification factor (F), four distinct uniplex PCR reactions were run at using primers 1117-70-17 and 1117-70-18 at concentrations of 0.01uM, 0.012 uM, 0.014 uM, 0.020 uM respectively. The four independent PCR reactions were carried out under the following conditions; 100mM KCl, 3mM MgCl, 10mM Tris pH 8.0, 200uM dNTPs using 10ng of hgDNA as template. Incubation was carried out at (94C/30 sec., 50C/20 sec.) for 30 cycles. Following PCR, reactions were diluted 1:10 with water and run under standard conditions using INVADER Assay FRET Detection Plates, 96 well genomic biplex, 100ng CLEAVASE VIII enzyme. Each 15ul reaction was set up as follows; 1ul of 1:10 diluted PCR reaction, 3ul of the PPI mix SNP#47932, 5ul 22.5mM MgCl2, 6ul of water, 15 ul of Chillout. The entire plate was incubated at 95C for 5min, and then at 63C for 60 min at which point a single read was taken on a Cytofluor 4000 fluorescent plate reader. For each of the four different primer concentrations (0.01uM, 0.012 uM, 0.014 uM, 0.020 uM) the amplification factor F was calculated using equation 1a, with $FOZm$=the sum of FOZ_FAM and FOZ_RED at 60 minutes, m=60, and FOZ_{240}=1.7. In plotting the primer concentration of each reaction against the log of the amplification factor Log(F), a strong linear relationship was noted (Figure 20). Using the data points in Figure 20, the formula describing

the linear relationship between amplification factor and primer concentration is described in equation 2:

$$Y=1.684X+2.6837 \text{ (equation 2a)}$$

Using equation 2, the amplification factor of a given amplicon Log(F)=Y could be manipulated in a predictable fashion using a known concentration of primer (X). In a converse manner, amplification bias observed under conditions of equimolar primer concentrations in multiplex PCR, could be measured as the "apparent" primer concentration (X) based on the amplification factor *F*. In multiplex PCR, values of "apparent" primer concentration among different amplicons can be used to estimate the amount of primer of each amplicon required to equalize amplification of different loci:

$$X=(Y-2.6837)/1.68 \text{ (equation 2b)}$$

Section 4.Calculation of Apparent Primer Concentrations from a Balanced Multiplex Mix.

As described in a previous section, primer concentration can directly influence the amplification factor of given amplicon. Under conditions of equimolar amounts of primers, *FOZm* readings can be used to calculate the "apparent" primer concentration of each amplicon using equation 2. Replacing Y in equation 2 with log(F) of a given amplification factor and solving for X, gives an "apparent" primer concentration based on the relative abundance of a given amplicon in a multiplex reaction. Using equation 2 to calculate the "apparent" primer concentration of all primers (provided in equimolar concentration) in a multiplex reaction, provides a means of normalizing primer sets against each other. In order to derive the relative amounts of each primer that should be added to an "Optimized" multiplex primer mix *R*, each of the "apparent" primer concentrations should be divided into the maximum apparent primer concentration (X_{max}), such that the strongest amplicon is set to a value of 1 and the remaining amplicons to values equal or greater than 1

$$R[n]=Xmax/X[n] \text{ (equation 3)}$$

Using the values of R[n] as an arbitrary value of relative primer concentration, the values of R[n] are multiplied by a constant primer concentration to provide working concentrations for each primer in a given multiplex reaction. In the example shown, the amplicon corresponding to

SNP assay 41646 has an R[n] value equal to 1. All of the R[n] values were multiplied by 0.01uM (the original starting primer concentration in the equimolar multiplex pcr reaction) such that lowest primer concentration is R[n] of 41646 which is set to 1, or 0.01uM. The remainder of the primer sets were also proportionally increased as shown in Figure 21. The results of multiplex PCR with the "optimized" primer mix are described below.

Section 5 Using optimized primer concentrations in multiplex PCR, variation in FOZ's among 10 INVADER assays are greatly reduced.

Multiplex PCR was carried out using 10-plex PCR using varying amounts of primer based on the volumes indicated in Figure 21 (X[max] was SNP41646, setting 1x=0.01uM/primer). Multiplex PCR was carried out under conditions identical to those used in with equimolar primer mix;100mMKCl, 3mMMgCl, 10mM Tris pH8.0, 200uM dNTPs, 2.5U taq, and 10ng of hgDNA template in a 50ul reaction. The reaction was incubated for (94C/30sec, 50C/44sec.) for 30 cycles. After incubation, the multiplex PCR reaction was diluted 1:10 with water and subjected to INVADER analysis. Using INVADER Assay FRET Detection Plates, (96 well genomic biplex, 100ng CLEAVASE VIII enzyme), reactions were assembled as 15ul reactions as follows; 1ul of the 1:10 dilution of the PCR reaction, 3ul of the appropriate PPI mix, 5ul of 22.5 mM MgCl2, 6ul of dH20. An additional 15ul of CHILL OUT was added to each well, followed by incubation at 95C for 5min. Plates were incubated at 63C and fluorescence measured on a Cytofluor 4000 at 10 min.

Using the following criteria to accurately make genotyping calls (FOZ_FAM+FOZ_RED-2 > 0.6), all 10 of 10 (100%) INVADER calls can be made after 10 minutes of incubation at 63C. In addition, the values of FAM+RED-2 (an indicator of overall signal generation, directly related to amplification factor (see equation 2)) varied by less than seven fold between the lowest signal (Figure 22, 67325, FAM+RED-2=0.7) and the highest (Figure 22, 47892, FAM+RED-2=4.3).

ii. PCR Primer Design Example 2

Using the TWT Oligo Order Entry Database, 144 sequences of less than 200 nucleotides in length were obtained with SNP annotated using brackets to indicate the SNP position for each

sequence (e.g. NNNNNNN[$N_{(wt)}$/$N_{(mt)}$]NNNNNNNN). In order to expand sequence data flanking the SNP of interest, sequences were expanded to approximately 1kB in length (500 nts flanking each side of the SNP) using BLAST analysis. Of the 144 starting sequences, 16 could not expanded by BLAST, resulting in a final set of 128 sequences expanded to approximately 1kB length (See, Figure 23). These expanded sequences were provided to the user in Excel format with the following information for each sequence; (1) TWT Number, (2) Short Name Identifier, and (3) sequence (see Figure 23). The Excel file was converted to a comma delimited format and used as the input file for Primer Designer INVADER CREATOR v1.3.3. software (this version of the program does not screen for FRET reactivity of the primers, nor does it allow the user to specify the maximum length of the primer). INVADER CREATOR Primer Designer v1.3.3., was run using default conditions (e.g. minimum primer size of 12, maximum of 30), with the exception of Tm_{low} which was set to 60C. The output file (see Figure 24, bottom of each sheet shows footprint region in upper case letters and SNP in brackets) contained 128 primer sets (256 primers, See Figure 25), four of which were thrown out due to excessively long primer sequences (SNP # 47854, 47889, 54874, 67396), leaving 124 primers sets (248 primers) available for synthesis. The remaining primers were synthesized using standard procedures at the 200nmol scale and purified by desalting. After synthesis failures, 107 primer sets were available for assembly of an equimolar 107-plex primer mix (214 primers, See Figure 25). Of the 107 primer sets available for amplification, only 101 were present on the INVADER MAP plate to evaluate amplification factor.

Multiplex PCR was carried out using 101-plex PCR using equimolar amounts of primer (0.025uM/primer) under the following conditions;100mMKCl, 3mM MgCl, 10mM Tris pH8.0, 200uM dNTPs, and 10ng of human genomic DNA (hgDNA) template in a 50ul reaction. After denaturation at 95C for 10min, 2.5 units of Taq was added and the reaction incubated for (94C/30sec, 50C/44sec.) for 50 cycles. After incubation, the multiplex PCR reaction was diluted 1:24 with water and subjected to INVADER assay analysis using INVADER MAP detection platform. Each INVADER MAP assay was run as a 6ul reaction as follows; 3ul of the 1:24 dilution of the PCR reaction (total dilution 1:8 equaling D=0.125), 3ul of 15 mM MgCl2 covered with covered with 6ul of CHILLOUT. Samples were denatured in the INVADER MAP plate by incubation at 95C for 5min., followed by incubation at 63C and fluorescence measured on a

Cytofluor 4000 (384 well reader) at various timepoints over 160 minutes. Analysis of the FOZ values calculated at 10, 20, 40, 80, 160 min. shows that correct calls (compared to genomic calls of the same DNA sample) could be made for 94 of the 101 amplicons detectable by the INVADER MAP platform (figure 26 and figure 27). This provides proof that the INVADER CREATOR Primer Designer software can create primer sets which function in highly multiplex PCR.

In using the FOZ values obtained throughout the 160 min. time course, amplification factor F and R[n] were calculated for each of the 101 amplicons (Figure 28). R[nmax] was set at 1.6, which although Low end corrections were made for amplicons which failed to provide sufficient FOZm signal at 160 min., assigning an arbitrary value of 12 for R[n]. High end corrections for amplicons whose FOZm values at the 10 min. read, an R[n] value of 1 was arbitrarily assigned. Optimized primer concentrations of the 101-plex were calculated using the basic principles outlined in the 10-plex example and equation 1b, with an R[n] of 1 corresponding to 0.025uM primer (see Fig.15 for various primer concentrations). Multiplex PCR was under the following conditions;100mMKCl, 3mM MgCl, 10mM Tris pH8.0, 200uM dNTPs, and 10ng of human genomic DNA (hgDNA) template in a 50ul reaction. After denaturation at 95C for 10min, 2.5 units of Taq was added and the reaction incubated for (94C/30sec, 50C/44sec.) for 50 cycles. After incubation, the multiplex PCR reaction was diluted 1:24 with water and subjected to INVADER analysis using INVADER MAP detection platform. Each INVADER MAP assay was run as a 6ul reaction as follows; 3ul of the 1:24 dilution of the PCR reaction (total dilution 1:8 equaling *D*=0.125), 3ul of 15 mM MgCl2 covered with covered with 6ul of CHILLOUT. Samples were denatured in the INVADER MAP plate by incubation at 95C for 5min., followed by incubation at 63C and fluorescence measured on a Cytofluor 4000 (384 well reader) at various timepoints over 160 minutes. Analysis of the FOZ values was carried out at 10, 20, and 40 min. and compared to calls made directly against the genomic DNA. Shown in Figure 26, is a comparison between calls made at 10 min. with a 101-plex PCR with the equimolar primer concentrations versus calls that were made at 10 min. with a 101-plex PCR run under optimized primer concentrations. Additional data for this example is shown in figures 29a, 29b, and 30). Under equimolar primer concentration, multiplex PCR results in only 50 correct calls at the 10 min time point, where under optimized primer concentrations multiplex

PCR results in 71 correct calls, resulting in a gain of 21 (42%) new calls. Although all 101 calls could not be made at the 10 min timepoint, 94 calls could be made at the 40 min. timepoint suggesting the amplification efficiency of the majority of amplicons had improved. Unlike the 10-plex optimization that only required a single round of optimization, multiple rounds of optimization may be required for more complex multiplexing reactions to balance the amplification of all loci.

Additional primers for CYP2D6 are shown in figure 31. Figure 32 shows one protocol for multiplex optimization.

F. Sample Preparation Component Design

In some embodiments, genomic DNA that contains a target sequence to be analyzed by the detection assay is used as a starting material for the detection assay. In some such embodiments, it may be desirable to amplify the one or more regions of the genomic DNA (e.g., to generate a plurality of target sequences to be detected). The present invention is not limited by the nature of the amplification technology employed. Amplification techniques include, but are not limited to, PCR and the technologies disclosed in U.S. Pat. Nos. 6,345,514 and 6,221,635, as well as foreign patents and applications, EP1113082, WO200146463, WO200146462, JP2001149097, JP 2001136954, and JP2001008660, herein incorporated by reference in their entireties. In certain embodiments, Rubicon OmniPlex technology is employed for sample preparation. Rubicon OmniPlex technology (See e.g., U.S. Pat. No. 6,197,557, herein incorporated by reference in its entirety) reformats naturally occurring chromosomes into new molecules called Plexisomes. Plexisomes represent the complete genome as amplifiable DNA units of equal length that function as a molecular relational database from which the genetic information can be more quickly and accurately recovered. Use of the technology avoids PCR amplification for sample preparation and for genotyping and haplotyping for gene discovery, pharmacogenomics, and diagnostics by providing highly multiplexing and sample amplification. In preferred embodiments, all the various components for running any of these sample preparation methods are included in a kit (e.g. with at least a portion of a detection assay).

III. Detection Assay Production

The present invention provides a high-throughput detection assay production system, allowing for high-speed, efficient production of thousands of detection assays. The high-throughput production systems and methods allow sufficient production capacity to facilitate full implementation of the funnel process described above—allowing comprehensive of all known (and newly identified) markers. Figure 98 shows a general overview of the oligonucleotide production and processing systems of the present invention.

In some embodiments of the present invention, oligonucleotides and/or other detection assay components (*e.g.*, those designed by the INVADERCREATOR software and directed to target sequences analyzed by the in silico systems and methods) are synthesized. In preferred embodiments, oligonucleotide synthesis is performed in an automated and coordinated manner. As discussed in more detail below, in some embodiments, produced detection assay are tested against a plurality of samples representing two or more different individuals or alleles (e.g., samples containing sequences from individuals with different ethnic backgrounds, disease states, etc.) to demonstrate the viability of the assay with different individuals. In some embodiments, the systems of the present invention allow at least 300 detection assays to be produced per day. In other embodiments, the systems of the present invention allow at least 1000, or at least 2000 detection assays to be produced per day.

In some embodiments, the present invention provides an automated DNA production process. In some embodiments, the automated DNA production process includes an oligonucleotide synthesizer component and an oligonucleotide processing component. In some embodiments, the oligonucleotide production component includes multiple components, including but not limited to, an oligonucleotide cleavage and deprotection component, an oligonucleotide purification component, an oligonucleotide dry down component; an oligonucleotide de-salting component, an oligonucleotide dilute and fill component, and a quality control component. In some embodiments, the automated DNA production process of the present invention further includes automated design software and supporting computer terminals and connections, a product tracking system (*e.g.*, a bar code system), and a centralized packaging component. In some embodiments, the components are combined in an integrated, centrally controlled, automated production system. The present invention thus provides methods

of synthesizing several related oligonucleotides (*e.g.*, components of a kit) in a coordinated manner. The automated production systems of the present invention allow large-scale automated production of detection assays for numerous different target sequences.

In certain embodiments, detection assays are produced in an in-line fashion, such that the synthesized and processed oligonucleotides remain in the same columns and/same holder (e.g. 96 or 384 well plate). In this regard, human and machine interaction with the oligonucleotides being manufactured is minimized.

In certain embodiments, the various production components (e.g. oligonucleotide synthesis component and the various oligonucleotide processing components) are grouped at a single manufacturing location. In different embodiments, the various components are not grouped. For example, the Inventory Control component may be in one location (e.g. closer to a base of customers, or closer to a particular supplier) while the synthesis components are in another location, and many of the processing components are in a third location. This type of remote manufacturing is made possible, for example, by the data management systems of the present invention that allow product orders and inventory for individual assays, and individual components of assays to be tracked. Also, the production and processing facilities may be grouped for ease of use, but there may be multiple locations each producing a different component of an assay. Again, the data management systems of the present invention allow these assay components be separately tracked and assembled in finished assays.

A. Oligonucleotide Synthesis Component

Once a particular oligonucleotide sequence or set of sequences has been chosen, sequences are sent (*e.g.*, electronically) to a high-throughput oligonucleotide synthesizer component. In some preferred embodiments, the high-throughput synthesizer component contains multiple DNA synthesizers.

In some embodiments, the synthesizers are arranged in banks. For example, a given bank of synthesizers may be used to produce one set of oligonucleotides (*e.g.*, for an INVADER or PCR reaction). The present invention is not limited to any one synthesizer. Indeed, a variety of synthesizers are contemplated, including, but not limited to MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), OligoPilot (Amersham Pharmacia,), the 3900

and 3948 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA), POLYPLEX (Genemachines), 8909 EXPEDITE, Blue Hedgehog (Metabio), MerMade (BioAutomation, Plano, Texas), Polygen (Distribio, France), PrimerStation 960 (Intelligent Bio-Instruments, Cambridge, MA), and the high-throughput synthesizer described in PCT Publication WO 01/41918. In some embodiments, synthesizers are modified or are wholly fabricated to meet physical or performance specifications particularly preferred for use in the synthesis component of the present invention. In some embodiments, two or more different DNA synthesizers are combined in one bank in order to optimize the quantities of different oligonucleotides needed. This allows for the rapid synthesis (*e.g.*, in less than 4 hours) of an entire set of oligonucleotides (all the oligonucleotide components needed for a particular assay, *e.g.*, for detection of one SNP using an INVADER assay). In certain embodiments, the synthesizers are configured for generating oligonucleotides in 96 or 384 well plates.

In some embodiments the DNA synthesizer component includes at least 100 synthesizers. In other embodiments, the DNA synthesizer component includes at least 200 synthesizers. In still other embodiments, the DNA synthesizer component includes at least 250 synthesizers. In some embodiments, the DNA synthesizers are run 24 hours a day.

1. Synthesizers

A. Exemplary Synthesizers

The present invention provides nucleic acid synthesizers and methods of using and modifying nucleic acid synthesizers. For example, the present invention provides highly efficient, reliable, and safe synthesizers that find use, for example, in high throughput and automated nucleic acid synthesis (*e.g.* arrays of synthesizers), as well as methods of modifying pre-existing synthesizers to improve efficiency, reliability, and safety.

A problem with currently available synthesizers is the emission of undesirable gaseous or liquid materials that pose health, environmental, and explosive hazards. Such emissions result from both the normal operation of the instrument and from instrument failures. Emissions that result from instrument failures cause a reduction or loss of synthesis efficiency and can provoke further failures and/or complete synthesizer failure. Correction of failures may require taking the

synthesizer off-line for cleaning and repair. The present invention provides nucleic acid synthesizers with components that reduce or eliminate unwanted emissions and that compensate for and facilitate the removal of unwanted emissions, to the extent that they occur at all. The present invention also provides waste handling systems to eliminate or reduce exposure of emissions to the users or the environment. Such systems find use with individual synthesizers, as well as in large-scale synthesis facilities comprising many synthesizers (*e.g.* arrays of synthesizers).

In some particularly preferred embodiments, the present invention provides efficient and safe "open system synthesizers." Open system synthesizers are contrasted to "closed system synthesizers" in that the reagent delivery, synthesis compartments, and waste extraction for each synthesis column are not contained in a system that remains physically closed (*i.e.*, closed from both the ambient environment and from the other synthesis columns in the same instrument) for the duration of the synthesis run. For example, in a closed system, tubing (or other means) provided for the addition and removal of reagent to each reaction compartment or synthesis column is generally fixed to the column with a coupling that is sealed to isolate the contents of that system from its surroundings. In contrast, in an open system, the dispensing and/or removal of reagent may be through means that are not physically coupled to the reaction compartment.

Further, a common dispensing or waste removal means may be shared by multiple reaction compartments, such that each compartment sharing the means is serviced in turn. An example of an "open system synthesizer" is described in PCT Publication WO 99/65602, herein incorporated by reference in its entirety. This publication describes a rotary synthesizer for parallel synthesis of multiple oligonucleotides. The tubing that supplies the synthesis reagents to the synthesis column does not form a continuous closed seal to the synthesis columns. Instead, the rotor turns, exposing the synthesis columns, in series, to the dispense lines, which inject synthesis reagents into the synthesis column. Open synthesizers offer advantages over closed synthesizers for the simultaneous production of multiple oligonucleotides. For example, a large number of independent synthesis columns, each intended to produce a distinct oligonucleotide, are exposed to a smaller number of dedicated reagent dispensers (*e.g.*, four dedicated dispensers for each of the nucleotides). Open systems also provide easy access to synthesis columns, which

can be added or removed without detaching any otherwise fixed connections to reagent dispensing tubing.

While open synthesizers have advantages for the production of oligonucleotides, they suffer from increased problems of emissions and failures. The direct exposure of the columns to their surroundings and the non-continuous path of reagents increases the number of points at which gaseous and liquid emissions occur, thereby increasing the release of unwanted emissions to the atmosphere and leakage within the synthesizer. Many synthesizers carry out reagent delivery, nucleic acid synthesis, and waste disposal under pressurized conditions. Open systems have frequent problems with loss of pressure, resulting in instrument failures and/or loss of synthesis efficiency. The open system synthesizers of the present invention dramatically reduce instrument failures and the corresponding emissions.

Whether a system used is open or closed, oligonucleotide synthesis involves the use of an array of hazardous materials, including but not limited to methylene chloride, pyridine, acetic anhydride, 2,6-lutidine, acetonitrile, tetrahydrofurane, and toluene. These reagents can have a variety of harmful effects on those who may be exposed to them. They can be mildly or extremely irritating or toxic upon short-term exposure; several are more severely toxic and/or carcinogenic with long-term exposure. Many can create a fire or explosion hazard if not properly contained. In addition, many of these chemicals must be assessed for emissions from normal operations, *e.g* for determining compliance with OSHA or environmental agency standards. Malfunction of a system, *e.g.*, as recited above, increases such emissions, thereby increasing the risk of operator exposure, and increasing the risk that an instrument may need to be shut down until risk to an operator is reduced and until any regulatory requirements for operation are met.

Emission or leakage of reagents during operation can have consequences beyond risks to personnel and to the environment. As noted above, instruments may need to be removed from operation for cleaning, leading to a temporary decrease in production capacity of a synthesis facility. Further, any emission or leakage may cause damage to parts of the instrument or to other instruments or aspects of the facility, necessitating repair or replacement of any such parts or aspects, increasing the time and cost of bringing an instrument back into operation. Failure to address emissions or leakage concerns may lead to additional expenses for operation of a facility,

e.g., costs for increased or improved fire or explosion containment measures, and addition of costs associated with the elimination of any instrument systems or wiring that have not been determined to be safe for use in such hazardous locations (*e.g.*, by reference to controlling codes, such as electrical codes, or codes covering operations in the presence of flammable and combustible liquids).

The synthesizers of the present invention provide a number of novel features that dramatically improve synthesizer performance and safety compared to available synthesizers. These novel features work both independently and in conjunction to provide enhanced performance. For example, in some embodiments, the synthesizers of the present invention prevent loss of pressure during synthesis and waste disposal. By preventing loss of pressure, synthesis columns are purged properly and do not overflow during subsequent synthesis steps. Thus, prevention of pressure loss further prevents liquid overflow and instrument contamination. Additionally, in some embodiments, sufficient pressure differentials are maintained across all columns to allow efficient synthesis and purging without instrument failure. For example, regardless of whether synthesis columns are actively involved in a particular round of synthesis (*e.g.*, short oligonucleotides will be completed prior to the completion of longer oligonucleotides and will not be actively synthesized during the later round of synthesis), sufficient pressure differentials are maintained to allow reagent delivery and purging from the active columns. A number of additional features of the synthesizers of the present invention are described in detail below.

In addition to providing efficient synthesizers, the present invention provides methods for modifying existing synthesizers to improve their efficiency. For example, one or more of the novel components of the present invention may be added into or substituted into existing synthesizers to improve efficiency and performance.

The present invention further provides means of reducing exposure of operators and the environment to synthesis reagents and waste. In one embodiment, the present invention reduces exposure by improving collection and disposal of emissions that occur during the normal operation of various synthesis instruments. In another embodiment, the present invention reduces exposure by improving aspects of the instrument to reduce risk of malfunctions leading to reagent escape from the system, *e.g.*, through leakage, overflow or other spillage.

While the present invention will be described with reference to several specific embodiments, the description is illustrative of the present invention and is not to be construed as limiting the invention. Various modifications to the present invention can be made without departing from the scope and spirit of the present invention. For example, much of the following description is provided in the context of an open system synthesizer (*see, e.g.*, WO99/65602). However, the invention is not limited to open system synthesizers.

In preferred embodiments, the present invention provides open-system solid phase synthesizers that are suitable for use in large-scale polymer production facilities. Each synthesizer is itself capable of producing large volumes of polymers. However, the present invention provides systems for integrating multiple synthesizers into a production facility, to further increase production capabilities.

Figure 33 illustrates a synthesizer 1. The synthesizer 1 is designed for building a polymer chain by sequentially adding polymer units to a solid support in a liquid reagent. The liquid reagents used for synthesizing oligonucleotides may vary, as the successful operation of the present invention is not limited to any particular coupling chemistry. Examples of suitable liquid reagents include, but are not limited to: Acetonitrile (wash); 2.5% dichloroacetic acid in methylene chloride (deblock); 3% tetrazole in acetonitrile (activator); 2.5% cyanoethyl phosphoramidite in acetonitrile (A, C, G, T); 2.5% iodine in 9% water, 0.5% pyridine, 90.5% THF (oxidizer); 10% acetic anhydride in tetrahydrofuran (CAP A); and 10% 1-methylimidazole, 10% pyridine, 80% THF. Various useful reagents and coupling chemistries are described in U.S. Pat. 5,472,672 to Bennan, and U.S. Pat. No. 5,368,823 to McGraw *et al.* (both of which are herein incorporated by reference in their entireties).

The solid support generally resides within a synthesis column and various liquid reagents are sequentially added to the synthesis column. Before an additional liquid reagent is added to a synthesis column, the previous liquid reagent is preferably purged from the synthesis column. Although the synthesizer 1 is particularly suited for building nucleic acid sequences, the synthesizer 1 is also configured to build any other desired polymer chain or organic compound (*e.g.* peptide sequences).

The synthesizer 1 preferably comprises at least one bank of valves and at least

one bank of synthesis columns. Within each bank of synthesis columns, there is at least one synthesis column for holding the solid support and for containing a liquid reagent such that a polymer chain can be synthesized. Within the bank of valves, there are preferably a plurality of valves configured for selectively dispensing a liquid reagent into one of the synthesis columns. The synthesizer 1 is preferably configured to allow each bank of synthesis columns to be selectively purged of the presently held liquid reagent. In particularly preferred embodiments, the synthesizer of the present invention is configured to allow synthesis columns within a bank to be purged even when not all of the synthesis columns contain liquid reagents (*e.g.* only a portion of the synthesis columns in a bank received a liquid reagent (*i.e.* "active"), while the remaining synthesis columns are no longer receiving liquid reagent (*i.e.* "idle"). For example, in some preferred embodiments of the present invention, the design of the material in the synthesis columns allows idle columns to resist the downward pressure of gas, thus making this pressure available to purge the synthesis columns that contain liquid reagent. Additional banks of valves provide the synthesizer 1 with greater flexibility. For example, each bank of valves can be configured to distribute liquid reagents to a particular bank of synthesis columns in a parallel fashion to minimize the processing time.

Multiple banks of valves can also be configured to distribute liquid reagents to a particular bank of synthesis columns in series. This allows the synthesizer 1 to hold a larger number of different reagents, thus being able to create varied nucleic acid sequences (*e.g.* 48 oligonucleotides, each with a unique sequence).

Figure 33 illustrates a top view of a rotary synthesizer 1. As illustrated in Figure 33, the synthesizer 1 includes a base 2, a cartridge 3, a first bank of synthesis columns 4, a second bank of synthesis columns 5, a plurality of dispense lines 6, a plurality of fittings 7 (a first bank of fittings 13, and a second bank of fittings 14), a first bank of valves 8 and a second bank of valves 9. Within each of the banks of valves 8 and 9, there is preferably at least one valve. Within each of the banks of synthesis columns 4 and 5, there is preferably at least one synthesis column. Each of the valves is capable of selectively dispensing a liquid reagent into one of the synthesis columns. Each of the synthesis columns is preferably configured for retaining a solid support such as polystyrene or CPG and holding a liquid reagent. Further, as each liquid reagent is

sequentially deposited within the synthesis column and sequentially purged therefrom, a polymer chain is generated (*e.g.* nucleic acid sequence).

Preferably, there is a plurality of reservoirs, each containing a specific liquid reagent to be dispensed to one of the plurality of valves 8 or 9. Each of the valves within the first bank and second bank of valves 8 and 9, is coupled to a corresponding reservoir. Each of the plurality of reservoirs is pressurized (*e.g.* by argon gas). As a result, as each valve is opened, a particular liquid reagent from the corresponding reservoir is dispensed to a corresponding synthesis column. Each of the plurality of dispense lines 6 is coupled to a corresponding one of the valves within the first and second banks of valves 8 and 9. Each of the plurality of dispense lines 6 provides a conduit for transferring a liquid reagent from the valve to a corresponding synthesis column. Each one of the plurality of dispense lines 6 is preferably configured to be flexible and semi-resilient in nature. In preferred embodiments, the dispense lines of the present invention have a large bore size to prevent clogging. In preferred embodiments, the internal diameter of the dispense tube is at least 0.25mm. In other embodiments, the internal diameter of the tube is at least 0.50mm or at least 0.75mm. In some embodiments, the internal diameter of the tube is greater than or equal to 1.0mm (*e.g.* 1.0mm, or 1.2mm, or 1.4mm). Preferably, the plurality of dispense lines 6 are each made of a material such as PEEK, glass, or coated with TEFLON or Parlene, or coated/uncoated stainless steel or other metallic material. Of course other materials may also be used. For example, useful characteristics of the material used for the dispense lines would be resistance to degradation by the liquid reagents, minimal "wetting" by the liquid reagents, ease of fabrication, relative rigidity, and ability to be produced with a smooth surface finish. Metallic tubing (*e.g.* stainless steel), benefit from electropolishing to improve the surface finish (*e.g.* in coated or uncoated application). Another important characteristic of useful dispense lines in the ability to provide a seal between the plurality of valves 10 and the plurality of fittings 7.

Each of the plurality of fittings 7 is preferably coupled to one of the plurality of dispense lines 6. The plurality of fittings 7 are preferably configured to prevent the reagent from splashing outside the synthesis column as the reagent is dispensed from the fitting to a particular synthesis column positioned below the fitting. In preferred embodiments, the fitting includes a nozzle that prevents reagents from drying at the point fluid exits the nozzle (*e.g.* prevents dried

reagents from causing the reagents stream to dispense at angles away from the intended synthesis column). Construction techniques to achieve consistent flow at the discharge point of the liquid reagents is achieved by the use of high quality parts and construction. For example, clean square cuts (without burrs or shavings), or the use of a "drawn tip" (*i.e.*, a tip of reduced diameter at the discharge point). The use of a drawn tip, for example, reduces the wall thickness at the point of discharge, thus reducing the area of the tube wall cross section, providing a smooth transition from the larger portion of the tube (reducing flow resistance) and increases the likelihood of a clean separation of the discharged liquid reagent from the tip of the tube. This clean "snap" of the liquid reagent minimizes the retention of the discharged fluid at the tip, and thus minimizes subsequent build up of any solids (*e.g.* dried reagent). Additionally, if a sharp cut off of the fluid flow is obtained, the fluid front will actually reside within the confines of the tube after discharge of the desired volume. This minimizes surface evaporation and helps to maintain a clean orifice (*e.g.* prevent reagent from drying at the tip). Another example of a useful technique to prevent liquid reagent from drying at the discharge point is providing a sleeve or sheath over the dispense line to a point near the tip (dispense point). This sleeve or sheath is particularly useful when employed in conjunction with a relatively flexible dispense line.

As shown in Figure 33, the first and second banks of valves 8 and 9 each have thirteen valves. In Figure 33, the number of valves in each bank is merely for exemplary purposes (*e.g.* other numbers of valves may be employed, like 14, 15, 16, 17, etc.).

Each of the synthesis columns within the first bank of synthesis columns 4 and the second bank of synthesis columns 5 is presently shown resting in one of a plurality of receiving holes 11 within the cartridge 3. Preferably, each of the synthesis columns within the corresponding plurality of receiving holes 11 is positioned in a substantially vertical orientation. Each of the synthesis columns is configured to retain a solid support such as polystyrene or CPG and hold liquid reagent(s). In preferred embodiments, polystyrene is employed as the solid support. Alternatively, any other appropriate solid support can be used to support the polymer chain being synthesized.

During synthesizer operation, each of the valves selectively dispenses a liquid reagent through one of the plurality of dispense lines 6 and fittings 7. The first and second banks of valves 8 and 9 are preferably coupled to the base 2 of the synthesizer 1. The cartridge 3 which

contains the plurality of synthesis columns 12 rotates relative to the synthesizer 1 and relative to the first and second banks of valves 8 and 9. By rotating the cartridge 3, a particular synthesis column 12 is positioned under a specific valve such that the corresponding reagent from this specific valve is dispensed into this synthesis column. In preferred embodiments, the cartridge 3 has a home position that allows the synthesizer to be properly aligned before operation (such that the liquid reagent is properly dispensed into the synthesis columns). Further, the first and second banks of valves 8 and 9 are capable of simultaneously and independently dispensing liquid reagents into corresponding synthesis columns.

A cross sectional view of synthesizer 1 is depicted in Figure 34. As depicted in Figure 34, the synthesizer 1 includes the base 2, a set of valves 15, a motor 16, a gearbox 17, a chamber bowl 18, a drain plate 19, a drain 20, a cartridge 3, a bottom chamber seal 21, a motor connector 22, a waste tube system 23, a controller 24, and a clear window 25. The valves 15 are coupled to base 2 of the synthesizer 1 and are preferably positioned above the cartridge 3 around the outside edge of the base 2. This set of valves 15 preferably contains fifteen individual valves which each deliver a corresponding liquid reagent in a specified quantity to a synthesis column held in the cartridge 3 positioned below the valves. Each of the valves may dispense the same or different liquid reagents depending on the user-selected configuration. When more than one valve dispenses the same reagent, the set of valves 15 is capable of simultaneously dispensing a reagent to multiple synthesis columns within the cartridge 3. When the valves 15 each contain different reagents, each one of the valves 15 is capable of dispensing a corresponding liquid reagents to any one of the synthesis columns within the cartridge 3.

The synthesizer 1 may have multiple sets of valves. The plurality of valves within the multiple sets of valves may be configured in a variety of ways to dispense the liquid reagents to a select one or more of the synthesis columns. For example, in one configuration, where each set of valves is identically configured, the synthesizer 1 is capable of simultaneously dispensing the same reagent in parallel from multiple sets of valves to corresponding banks of synthesis columns. In this configuration, the multiple banks of synthesis columns may be processed in parallel. In the alternative, each individual valve within multiple sets of valves may contain entirely different liquid reagents such that there is no duplication of reagents among any individual valves in the multiple sets of valves. This configuration allows the synthesizer 1 to

build polymer chains requiring a large variety of reagents without changing the reagents associated with each valve.

The motor 16 is preferably mounted to the base 2 through the gear box 17 and the motor connector 22. The chamber bowl 18 preferably surrounds the motor connector 22 and remains stationary relative to the base 2.

The chamber bowl 18 is designed to hold any reagent spilled from the plurality of synthesis columns 12 during the purging process (or the dispensing process). Further, the chamber bowl 18 is configured with a tall shoulder to insure that spills are contained within the bowl 18. The bottom chamber seal 21 preferably provides a seal around the motor connector 22 in order to prevent the contents of the chamber bowl 18 from flowing into the gear box 17 (see Figure 34). The bottom chamber seal 21 is preferably composed of a flexible and resilient material such as TEFLON (or elastomer which conforms to any irregularities of the motor connector 22). Alternatively, the bottom chamber seal can be composed of any other appropriate material. In particularly preferred embodiments, the bottom chamber seal is composed of material that resists constant contact with liquid reagents (*e.g.*, TEFLON or Parlene). Additionally, the bottom chamber seal 21 may have frictionless properties that allow the motor connector 22 to rotate freely within the seal. For example, coating this flexible material with TEFLON helps to achieve a low coefficient of friction.

The clear window 25 is attached to (formed in) a top cover 30 of the synthesizer 1 and covers the area above the cartridge 3. The top cover 30 of synthesizer 1 seals the top part of the chamber (when in place), and opens up allowing an operator or maintenance person access to the interior of the synthesizer 1. The clear window 25 in top cover 30 allows the operator to observe the synthesizer 1 in operation while providing a pressure sealed environment within the interior of the synthesizer 1. As shown in Figure 34, there are a plurality of through holes 26 in the clear window 25 to allow the plurality of dispense lines 6 to extend through the clear plate 25 to dispense material into the synthesis columns located in cartridge 3.

The clear window 25 also includes a gas fitting 27 attached therethrough. The gas fitting 27 is coupled to a gas line 28. The gas line 28 preferably continuously emits a stream of inert gas (*e.g.* Argon) which flows into the synthesizer 1 through the gas fitting 27 and flushes out traces of air and water from the plurality of synthesis columns 12 within the synthesizer 1.

Providing the inert gas flow through the gas fitting 27 into the synthesizer 1 prevents the polymer chains being formed within the synthesis columns from being contaminated without requiring the plurality of synthesis columns 12 to be hermetically sealed and isolated from the outside environment.

Figure 35 shows the cartridge 3 in chamber bowl 18, with the top plate 30 removed, thus revealing the top chamber seal 31. Top chamber seal 31 is designed to provide a tight seal between top plate 30 and chamber bowl 18, such that inert gas applied through clear window 25 does not leak. If the top chamber seal 31 does not function properly, the inert gas leaks out (lowering the pressure in the chamber), thus causing the purge operation (that relies on the pressure on the inert gas) to fail. When the purge operation fails, un-purged columns quickly fill up and overflow. In some embodiments, a V-seal type top chamber seal is employed to prevent leakage of gas. In some embodiments, the hinges and latches on top plate 30 (not shown) are precisely machined to provide balanced forces on the top plate 30, such that the top plate 30 fits tightly over the chamber bowl.

Figure 36 illustrates a detailed view of a cartridge 3 for synthesizer 1. Preferably, the cartridge 3 is circular in shape such that it is capable of rotating in a circular path relative to the base 2 and the first and second banks of valves 8 and 9. The cartridge 3 has a plurality of receiving holes 11 on its upper surface around the peripheral edge of the cartridge 3. Each of the plurality of receiving holes 11 is configured to hold one of the synthesis columns 12. The plurality of receiving holes 11, as shown on the cartridge 3, is divided up among four banks. A bank 32 illustrates one of the four banks on the cartridge 3 and contains twelve receiving holes, wherein each receiving hole is configured to hold a synthesis column. An exemplary synthesis column 12 is shown being inserted into one of the plurality of receiving holes 11. The total number of receiving holes shown on the cartridge 3 includes forty-eight (48) receiving holes, divided into four banks of twelve receiving holes each. The number of receiving holes and the configuration of the banks of receiving holes is shown on the cartridge 3 for exemplary purposes only. Any appropriate number of receiving holes and banks of receiving holes can be included in the cartridge 3. Preferably, the receiving holes 11 within the cartridge each have a precise diameter for accepting the synthesis columns 12, which also each have a corresponding precise exterior surface 61 (see Figure 44) to provide a pressure-tight seal when the synthesis columns

12 are inserted into the receiving holes 11. In preferred embodiments, the synthesis column includes a column seal 65 (see Figure 44), such as a ring seal or a ball seal (*e.g.*, a flexible TEFLON ring that flexes on engagement of the synthesis column in the receiving hole 11). In other preferred embodiments, a seal, such as a ring seal, is provided above or in the receiving holes 11 (see, *e.g.*, Figure 44).

Figure 37 depicts an exemplary drain plate 19 of the synthesizer 1. The drain plate 19 is coupled to the motor connector 22 (not shown) through securing holes 33. More specifically, the drain plate 19 is attached to the motor connector 22, which rotates the drain plate 19 while the motor 16 is operating and the gear box 17 is turning. The cartridge 3 and the drain plate 19 are preferably configured to rotate as a single unit. The drain plate 19 is configured to catch and direct the liquid reagents as the liquid reagents are expelled from the plurality of synthesis columns (during the purging process). During operation, the motor 16 is configured to rotate both the cartridge 3 and the drain plate 19 through the gear box 17 and the motor connector 22. The bottom chamber seal 21 allows the motor connector 22 to rotate the cartridge 3 and the drain plate 19 through a portion of the chamber bowl 18 while still containing spilled reagents in the chamber bowl 18. The controller 24 is coupled to the motor 16 to activate and deactivate the motor 16 in order to rotate the cartridge 3 and the drain plate 19. The controller 24 (see Figure 34) provides embedded control to the synthesizer and controls not only the operation of the motor 16, but also the operation of the valves 15 and the waste tube system 23.

The drain plate 19 has a plurality of securing holes 33 for attaching to the motor connector 22. The drain plate 19 also has a top surface 34 which may, in some embodiments, attach to the underside of the cartridge 3. In other embodiments, a drain plate gasket is provided between the drain plate 19 and cartridge 3 (see below). As stated previously, the cartridge 3 holds the plurality of synthesis columns grouped into a plurality of banks. The drain plate preferably has a collection area corresponding to each of the banks of synthesis columns (*e.g.* four in Figure 37 to correspond to the four banks of synthesis columns in cartridge 3). Each of these four collection areas 35, 36, 37 and 38 in Figure 37, forms a recessed area below the top surface 34 and is designed to contain and direct material flushed from the synthesis columns within the bank above the collection area.

Each of the four collection areas 35, 36, 37 and 38 is positioned below a corresponding one of the banks of synthesis columns on the cartridge 3. The drain plate 19 is rotated with the cartridge 3 to keep the corresponding collection area below the corresponding bank.

In Figure 37, there are four drains 39, 40, 41, and 42 each of which is located within one of the four collection areas 35, 36, 37 and 38 respectively. In use, the collection areas are configured to contain material flushed from corresponding synthesis columns and pass that material through the drains. Preferably, there is a collection area and a drain corresponding to each bank of synthesis columns within the cartridge 3. Alternatively, any appropriate number of collection areas and drains can be included within a drain plate. Figure 38A shows a top view of drain plate gaskets 43. The drain plate gasket is configured to be situated between drain plate 19 and cartridge 3. Drain plate gasket 43 is shown in Figure 38A with guide holes 44 and drain cut-outs 57, 58, 59, and 60. Guide holes 44 allow the drain plate gasket to fit over the motor connector 22. Drain cut-outs 57-60 allow the bottom column opening of synthesis columns 12 to discharge material into collection areas 35-38 in drain plate 19. In other embodiments, the drain cut outs mirror the receiving holes in the cartridge (see cut-outs 60 in Figure 38B), such that each column is able to discharge material into collection areas 35-38, while having a seal around each synthesis column. In some embodiments, all of the cut-outs are for the synthesis columns, like the cuts 60 depicted in Figure 38B.

The drain plate gaskets of the present invention may be made of any suitable material (*e.g.* that will provide a tight seal above drain plate 19, such that gas and liquid do not escape). In some embodiments, the drain plate gasket is composed of rubber. Providing a tight seal between cartridge 3 and drain plate 19 with a drain plate gasket helps maintain the proper pressure of inert gas during purging procedures, such that synthesis columns with liquid reagent properly drain (preventing overflow during the next cycle). The seal between cartridge 3 and drain plate 19 may also be improved by the addition of grease between the components, or very finely machining the contact points between the two components. In other embodiments, the seal between the cartridge and drain plate is improved by physically bonding the plates together, or machining either the cartridge or drain plate such that concentric ring seals may inserted into the machined component. In still other embodiments, the two components are manufactured as a single component (*e.g.* a single components with all the features of both the cartridge and drain

plate formed therein). In preferred embodiments, one component is provided with plurality of concentric circular rings that contact the flat surface of the other component and act as seals.

Figure 39 shows a side view of a drain plate gasket 43 situated between cartridge 3 and drain plate 19. Figure 39 also shows a drain 20 extending from drain plate 19. Figure 39 also shows a drain with sealing ring 45 (sealing ring is labeled 46). The sealing ring 46 tightly seals the connection between the drain 45 and the waste tube system 23 (see Figure 40). Also shown in Figure 39 is a synthesis column 12 inserted in cartridge 3, passing through drain plate gasket 43, and ending in drain plate 19.

The waste tube system 23 is preferably utilized to provide a pressurized environment for flushing material including reagents from the plurality of synthesis columns located within a corresponding bank of synthesis columns and expelling this material from the synthesizer 1. Alternatively, the waste tube system 23 can be used to provide a vacuum for drawing material from the plurality of synthesis columns located within a corresponding bank of synthesis columns.

A cross-sectional view of the waste tube system 23 is illustrated in Figure 39. The waste tube system 23 comprises a stationary tube 47 and a mobile waste tube 48. The stationary tube 47 and the mobile waste tube 48 are slidably coupled together. The stationary tube 47 is attached to the chamber bowl 18 and does not move relative to the chamber bowl (see Figure 41). In contrast, the mobile tube 48 is capable of sliding relative to the stationary tube 47 and the chamber bowl 18. When in an inactive state, the waste tube system 47 does not expel any reagents. During the inactive state, both the stationary tube 47 and the mobile tube 48 are preferably mounted flush with the bottom portion of the chamber bowl 18 (see Figure 41). When in an active state, the waste tube system 23 purges the material from the corresponding bank of synthesis columns. During the active state, the mobile tube 48 rises above the bottom portion of the chamber bowl 18 towards the drain plate 19. The drain plate 19 is rotated over to position a drain corresponding to the bank to be flushed, above the waste tube system 23. The mobile tube 48 then couples to the drain (*e.g.*, 20 or 45) and the material is flushed out of the corresponding bank of synthesis columns and into the drain plate 19. The liquid reagent is purged from the corresponding bank of synthesis columns due to a sufficient pressure differential between a top opening 49 (Figure 44) and a bottom opening 50 (Figure 44) of each synthesis

column. This sufficient pressure differential is preferably created by coupling the mobile waste tube 48 to the corresponding drain. Alternatively, the waste tube system 23 may also include a vacuum device 29 (see, Figure 34) coupled to the stationary tube 47 (see Figure 40) wherein the vacuum device 29 is configured to provide this sufficient pressure differential to expel material from the corresponding bank of synthesis columns. When this sufficient pressure differential is generated, the excess material within the synthesis columns being flushed, then flows through the corresponding drain and is carried away via the waste tube system 23.

When engaging the corresponding drain to flush a bank of synthesis columns, preferably the mobile tube 48 slides over the corresponding drain such that the mobile tube 48 and the drain act as a single unit. Alternatively, the waste tube system 23 includes a mobile tube 48 which engages the corresponding drain by positioning itself directly below the drain and then sealing against the drain without sliding over the drain. The mobile tube 48 may include a drain seal positioned on top of the mobile tube. In this embodiment, during a flushing operation, the mobile tube 48 is not locked to the corresponding drain. In the event that this drain is accidentally rotated while the mobile waste tube 48 is engaged with the drain, the drain and mobile tube 48 of the synthesizer 1 will simply disengage and will not be damaged. If this occurs while material is being flushed from a bank of synthesis columns, any spillage from the drain is contained within the chamber bowl 18. In preferred embodiments, the bottom of the chamber bowl 18 has a chamber drain 64 (see Figure 41) to collect and remove any spilled material in the chamber bowl. In this regard, material may be removed before it builds up and leaks into other parts of the synthesizer (*e.g.* motor 16 or gear box 17). In some embodiments of the present invention, the chamber drain is in a closed position during synthesis and purging. When the top cover of the synthesizer is opened, the chamber drain can be opened, drawing out unwanted gaseous or liquid emissions (*e.g.*, using a vacuum source). Coordination of the chamber drain opening to the top cover opening may be accomplished by mechanical or electric means.

Configuring the waste tube system 23 to expel the reagent while the mobile waste tube 48 is coupled to the drain allows the present invention to selectively purge individual banks of synthesis columns. Instead of simultaneously purging all the synthesis columns within the synthesizer 1, the present invention selectively purges individual banks of synthesis columns

such that only the synthesis columns within a selected bank or banks are purged. In preferred embodiments, the waste system is fitted for qualitative monitoring of detritylation. For example, colorimetric analysis of waste effluent using, for example, a CCD camera or a similar device provides a yes/no answer on a particular detritylation level. Qualitative analysis can also be accomplished by spectrophotometricly, or by testing effluent conductivity. Qualitative detection of detritylation can generally be performed with less expensive equipment than is generally required by more precise quantitation, and yet generally provides sufficient monitoring for detritylation failure. In preferred embodiments, the effluent from each column is monitored when a bank of columns is purged.

Preferably, the synthesizer 1 includes two waste tube systems 23 for flushing two banks of synthesis columns simultaneously. Alternatively, any appropriate number of waste tube systems can be included within the synthesizer 1 for selectively flushing synthesis columns or banks of synthesis columns. In preferred embodiments, the waste tube systems 23 are spaced on opposite sides of the chamber bowl 18 (*i.e.* they are directly across from each other, see Figure 41). In this regard, the force on the drain plate 19 is equalized during flushing procedures (*e.g.* the drain plate is less likely to tip one way or the other from force being applied to just one side of the plate). Alternatively, a single waste tube system 23 may be provided for flushing the plurality of banks of synthesis columns. When a single waste tube system is used, it is preferred that a balancing force be provided on the opposite side of the drain plate 19, *e.g.*, such as would be provided by the presence of a second waste tube system 23. In one embodiment, a balancing force is provided by a dummy waste tube system (not shown), that may be actuated in the same fashion as the waste tube system 23, but which does not serve to drain the bank of synthesis columns to which it is deployed.

In use, the controller 24, which is coupled to the motor 16, the valves 15, and the waste tube system 23, coordinates the operation of the synthesizer 1. The controller 24 controls the motor 16 such that the cartridge is rotated to align the correct synthesis columns with the dispense lines 6 corresponding to the appropriate valves 15 during dispensing operations and that the correct one of the drains 39, 40, 41, and 42 are aligned with an appropriate waste tube system 23 during a flushing operation.

In some preferred embodiments, the synthesizer comprises a means of delivering energy to the synthesis columns to, for example, increase nucleic acid coupling reaction speed and efficiency, allowing increased production capacity. In some embodiments, the delivery of energy comprises delivering heat to the chamber or the columns. In addition to increasing production capacity, the use of heat allows the use of alternate synthesis chemistries and methods, *e.g.*, the phosphate triester method, which has the advantages of using more stable monomer reagents for synthesis, and of not using tetrazole or its derivatives as condensation catalysts. Heat may be provided by a number of means, including, but not limited to, resistance heaters, visible or infrared light, microwaves, Peltier devices, transfer from fluids or gasses (*e.g.*, via channels or a jacketed system). In some embodiments, heat generated by another component of a synthesis or production facility system (*e.g.*, during a waste neutralization step) is used to provide heat to the chamber or the columns. In other embodiments, heat is delivered through the use of one or more heated reagents. Delivery of heat also comprises embodiments wherein heat is created within the, *e.g.*, by magnetic induction or microwave treatment. In some embodiments, heat is created at or within synthesis columns. It is contemplated that heating may be accomplished through a combination of two or more different means.

In some embodiments, the delivery of heat provides substantially uniform heating to two or more synthesis columns. In some embodiments, heating is carried out at a temperature in a range of about 20 °C to about 60 °C. The present invention also provides methods for determining an optimum temperature for a particular coupling chemistry. For example, multiple synthesizers are run side-by-side with each machine run at a different temperature. Coupling efficiencies are measured and the optimum temperature for one or more incubations times are determined. In other embodiments, different amounts of heat are delivered to different synthesis columns within a single synthesizer, such that different reaction chemistries or protocols can be run at the same time.

Delivery of heat to an enclosed, sealed system will alter the pressure within the system. It is contemplated that the sealed system of the present invention will be configured to tolerate variations in the system pressure (i.e., the pressure within the sealed system) related to heating or other energy input to the system. In preferred embodiments, the system (*e.g.*, every component of the system and every junction or seal within the system) will be configured to withstand a

range of pressures, *e.g.*, pressures ranging from 0 to at least 1 atm, or about 15 psi. It is contemplated that pressures may be varied between different points within the system. For example, in some embodiments, reagents and waste fluids are moved through the synthesis column by use of a pressure differential between one end (*e.g.*, an input aperture) and the other (*e.g.*, a drain aperture) of the synthesis column. In some embodiments, the system of the present invention is configured to use pressure differentials within a pressurized system (*e.g.*, wherein a system segment having lower pressure than another system segment nonetheless has higher pressure than the environment outside the sealed system). In some embodiments, the prevention of backward flow of reagents through the system (*e.g.*, in the event of back pressure from a process step such as heating) is controlled by use of pressure. In other embodiments, valves are provided to assist in control of the direction of flow.

In other preferred embodiments, the synthesizer comprises a mixing component configured to mix reaction components, *e.g.*, to facilitate the penetration of reagents into the pores of the solid support. Mixing may be accomplished in a number of ways. In some embodiments, mixing is accomplished by forced movement of the fluid through the matrix (*e.g.*, moving it back and forth or circulating it through the matrix using pressure and/or vacuum, or with a fluid oscillator). Mixing may also be accomplished by agitating the contents of the synthesis column (*e.g.*, stirring, shaking, continuous or pulsed ultra or subsonic waves). Examples are provided in Figures 42A-C, which illustrate different embodiments of energy input components 95 and mixing components 96. Also, Figures 43A-B illustrate different combinations of energy input components 95 and mixing components 96.

In some preferred embodiments, an agitator is used that avoids the creation of standing waves in the reaction mixture. In some preferred embodiments, the agitator is configured to utilize a reaction vessel surface or reaction support surface (*e.g.*, a surface of a synthesis column) to serve as resonant members to transfer energy into fluid within a reaction mixture. In a preferred embodiment, a horn is applied directly to the cartridge 3 to provided pulsed or continuous ultra sonic energy to the synthesis columns therein. In some embodiments, the matrix is an active component of the mixing system. For example, in some embodiments, the matrix comprises paramagnetic particles that may be moved through the use of magnets to facilitate mixing. In some embodiments, the matrix is an active component of both mixing and

heating systems (*e.g.*, paramagnetic particles may be agitated by magnetic control and heated by magnetic induction). It is contemplated that any of these mixing means may be used as the sole means of mixing, or that these mixing components may be used in combination, either simultaneously or in sequence. In preferred embodiments, the heating component and the mixing component are under automated control.

Figure 42 illustrates a cross sectional view of a synthesis column 12. The synthesis column is an integral portion of the synthesizer 1. Generally, the polymer chain is formed within the synthesis column 12. More specifically, the synthesis column 12 holds a solid support 54 on which the polymer chain is grown. Examples of suitable solid supports include, but are not limited to, polystyrene, controlled pore glass, and silica glass. As stated previously, to create the polymer chain, the solid support 54 is sequentially submerged in various reagents for a predetermined amount of time. With each deposit of a reagent, an additional unit is added, or the solid support is washed, or failure sequences are capped, etc. Preferably, the solid support 54 is held within the synthesis column 12 by a bottom frit 55. In particularly preferred embodiments, a top frit 53 is included above the solid support (*e.g.* to help resist downward gas pressure when the particular synthesis column does not have liquid reagents, but other synthesis columns within the bank are being purged of their liquid contents). The synthesis column 12 includes a top opening 49 and a bottom opening 50. During the dispensing process, the synthesis column 12 is filled with a reagent through the top opening 49. During the purging process, the synthesis column 12 is drained of the reagent through the bottom opening 50. The bottom frit 55 prevents the solid support from being flushed away during the purging process.

The exterior surface 61 of each synthesis column 12 fits within the receiving hole 11 within the cartridge 3 and provides a pressure tight seal around each synthesis column within the cartridge 3. Preferably, each synthesis column is formed of polyethylene or other suitable material. In preferred embodiments, the receiving holes 11 of the cartridge 3 are provided with seals, such as O-ring seals 67, that will flex on engagement of the synthesis column 12 in receiving hole 11 and accommodate any irregularities in the exterior surface 61 of the synthesis column 12, thus assuring the presence of a pressure-tight seal.

In preferred embodiments, the material inside the synthesis column (*e.g.* in Figure 44, this includes top frit 53, solid support 54, and bottom frit 55) is configured to resist the

downward pressure of gas (*e.g.*, to provide back pressure) applied during the purging process when the particular synthesis column does not have liquid reagent. In this regard, other synthesis columns that do contain liquid reagents may be successfully purged with the application of gas pressure during the purging process (*i.e.* the synthesis columns without liquid reagent do not allow a substantial portion the gas pressure applied during the purging process to escape through their bottom openings). Other packing materials may also be added to the synthesis columns to help maintain the pressure differential across the column when it is idle.

One method for constructing a synthesis column that successfully resists the downward pressure of gas (when no liquid reagent has been added to this column) is to include a top frit in addition to a bottom frit. Determining what type of top frit is suitable for any given synthesis column and type of solid support may be determined by test runs in the synthesizer. For example, the columns may be loaded into the synthesizer with the candidate top frit (and solid support and bottom frit), and instructions for synthesizing different length oligonucleotides inputted (i.e., this will allow certain columns to sit idle while other columns are still having liquid dispensed into them and purged out). Observation through the glass panel, examining the amount of leakage from overflowing columns, and testing the quality of the resulting oligonucleotides, are all methods to determine if the top frit is suitable (*e.g.*, a thicker or smaller pore top frit may be employed if problems associated with insufficient back pressure are seen). By combining the appropriate packing material in columns with the appropriate delivered pressure to the chamber, purging can be efficiently carried out, avoiding spill-over that can result in synthesis or instrument failure.

Another method for constructing a synthesis column that successfully resists the downward pressure of gas (when no liquid reagent has been added to this column) is to provide a solid support that resists this downward force even when no liquid reagent is in the columns. One suitable solid support material is polystyrene (*e.g.* US Pat. No. 5,935,527 to Andrus *et al.*, hereby incorporated by reference). In some embodiments, the styrene (of the polystyrene) is cross-linked with a cross-linking material (*e.g.* divinylbenzene). In some embodiments, the cross-linking ratio is 10-60 percent. In preferred embodiments, the cross-linking ration is 20-50 percent. In particularly preferred embodiments, the cross-linking ratio is about 30-50 percent. In some embodiments, the polystyrene solid support is used in conjunction with a top frit in order to

successfully resist the downward pressure of gas during the purging process. In some embodiments, the polystyrene is used as the solid support for synthesis. In other embodiments, a different support, such as controlled pore glass, is used as the support for the synthesis reaction, and the polystyrene is provided only to increase the back pressure from a column comprising a CPG or other synthesis support.

There are many advantages of configuring synthesis columns to successfully resist downward gas pressure during the purging process. One advantage is the fact that not all the synthesis columns need to contain liquid reagent during the purging process in order for the purge to be successful. Instead, one or more of the synthesis columns may remain idle during a particular cycle, while the other synthesis columns continue to receive liquid reagents. In this regard, oligonucleotides of different lengths may be constructed (*e.g.*, a 20-mer constructed in one synthesis column may be completed and sit idle, while a 32-mer is constructed in a second synthesis column). Achieving successful purges after each liquid addition prevents liquid leakage (*e.g.* additional liquid reagent applied to a synthesis column that was not successfully purged will cause the column to overflow).

Figure 45 illustrates a computer system 62 coupled to the synthesizer 11. The computer system 62 preferably provides the synthesizer 1, and specifically the controller 24, with operating instructions. These operating instructions may include, for example, rotating the cartridge 3 to a predetermined position, dispensing one of a plurality of reagents into selected synthesis columns through the valves 15 and dispense lines 6, flushing the first bank of synthesis columns 4 and/or the second bank of synthesis columns 5, and coordinating a timing sequence of these synthesizer functions. U.S. Patent 5,865,224 to Ally *et al.* (herein incorporated by reference in its entirety), further demonstrates computer control of synthesis machines. Preferably, the computer system 62 allows a user to input data representing oligonucleotide sequences to form a polymer chain via a graphical user interface.

After a user inputs this data, the computer system 62 instructs the synthesizer 1 to perform appropriate functions without any further input from the user. The computer system 62 preferably includes a processor, an input device and a display. The computer 62 can be configured as a laptop or a desktop, and may be operably connected to a network (*e.g.* LAN, internet, etc.).

In some embodiments, the present invention provides alignment detectors for detecting the alignment of any of the components of the present invention, as desired. In some embodiments, when a misalignment is detected, an alarm or other signal is provided so that a user can assure proper alignment prior to further operation. In other embodiments, when a misalignment is detected, a processor operates a motor to adjust that alignment. Alignment detectors find particular use in the present invention for assuring the alignment of any components that are involved in an exchange of liquid materials. For example, alignment of dispense lines and synthesis columns and alignment of drains and waste tubes should be monitored. Likewise, the tilt angle of the cartridge or any other component that should be parallel to the work surface can be monitored with alignment detectors.

As noted above, the exterior surface 61 of each synthesis column 12 fits within the receiving hole 11 within the cartridge 3 and is intended to provide a pressure-tight seal around each synthesis column 12 within the cartridge 3. Figure 46 illustrates three cross-sectional detailed views of the assembly 66 (the assembly comprising the cartridge 3, the drain plate gasket 43 and the drain plate 19) with a synthesis column 12 within a receiving hole 11 of cartridge 3. Each view shows a different embodiment of an airtight seal between the assembly 66 and the exterior surface 61 of synthesis column 12. In some embodiments, the airtight seal is provided by an O-ring 67. In preferred embodiments, the O-ring 67 is accessible for easy insertion and removal, *e.g.*, for cleaning or replacement. In one embodiment, an O-ring 67 is positioned at the top of receiving hole 11, held in place by, *e.g.*, a restraining plate 68, or any other suitable restraining fitting. In a preferred embodiment, a channel 69 is provided at the top of receiving hole 11 in cartridge 3 to accommodate the O-ring 67, as illustrated in Figure 46A. In a particularly preferred embodiment, a groove 70 within receiving hole 11 in cartridge 3 accommodates an O-ring 67, providing a groove lip 71 to restrain the O-ring 67, as illustrated in Figure 46B. In a particularly preferred embodiment, the groove lip 71 is about 0.030 inches. Figure 46C illustrates a further embodiment, in which drain plate gasket 43 is configured to provide an airtight seal between nucleic acid synthesis column 12 and assembly 66. The illustrations in Figure 46 are provided by way of examples only, and it is not intended that the present invention be limited by details of these illustrations, such as apparent size, shape or

precise locations of features such as grooves, channels, plates or seals. Any O-ring configuration that helps maintain proper pressure differential across the synthesis columns is contemplated.

O-rings 67 may be composed of any suitable material, preferably a chemically resistant, resilient material that flexes upon engagement of the synthesis column 12 in receiving hole 11. In some embodiments, a low cost material such as silicone or VITON may be used. In other embodiments, more expensive materials offering longer term stability, such as KALREZ, may be used. In some embodiments the O-rings may have a light lubrication, *e.g.* with a silicone or fluorinated grease.

In some embodiments, the present invention provides a means of collecting emissions from reagent reservoirs 72 (See *e.g.*, Figure 47A and B) by providing a reagent dispensing station. In one embodiment, the reagent dispensing station is an integral part of the base 2 of the synthesizer, as illustrated in Figures 47A and 47B. In some embodiments, the reagent dispensing station provides an enclosure for collecting emitted gasses. In some embodiments, the enclosure is created by the provision of a panel 73 to enclose a portion of base 2 containing reagent reservoirs 72, as illustrated in Figure 47B. In some embodiments, the panel 73 is movable for easy access to reagent reservoirs. In some embodiments, it is removeably attached. Removable attachment may be accomplished by any suitable means, such as through the use of VELCRO, screws, bolts, pins, magnets, temporary adhesives, and the like. In preferred embodiments, at least a portion of the panel 73 is slidably moveable. In preferred embodiments, at least a portion of panel 73 is transparent. In some embodiments, the enclosure of the reagent dispensing station comprises a viewing window that is not in a panel 73.

In some embodiments, the enclosure comprises a ventilation tube. In preferred embodiments, panel 73 comprises a ventilation port 74, *e.g.*, for attachment to a ventilation tube. Since reagent vapors are typically heavier than air, in preferred embodiments, the ventilation tube is attached at the bottom for the enclosure. In a particularly preferred embodiment, the ventilation port is positioned toward the rear of the instrument.

In some embodiments, the enclosure further comprises an air inlet. In a preferred embodiment, a clearance 75 between the panel 73 and the base 2 provides an air inlet. In a particularly preferred embodiment, the air inlet is positioned toward the front of the instrument.

The location of the ventilation port 74 and air inlet is not limited to the panel 73. For example, in an alternative embodiment, the reagent dispensing station comprises a stand for holding the reagent bottles and a ventilation tube, wherein the stand holds the reagent reservoirs and the ventilation tube removes emitted gases.

Ventilation may be continuous or under the control of an operator. For example, in some embodiments, when the panel 73 is in a closed position, ventilation occurs continuously through the ventilation port 74 or at regular intervals. In other embodiments, an operator may manually activate ventilation prior to opening the panel 73. In still other embodiments, ventilation occurs in an automated fashion immediately prior to the opening of panel 73. For example, where the opening of panel 73 is controlled by a computer processor, activation of the "open" routine triggers ventilation prior to the physical opening of panel 73. In still other embodiments, the contents of the reagent containers are monitored by a sensor and the ventilation is triggered when one or more of the reagent containers are depleted. In some embodiments, the panel 73 is also automatically open, indicating the need for additional reagents and/or allowing an automated reagent container delivery system to supply reagents to the system.

The present invention also provides systems for ventilation, particularly ventilation of reaction enclosures (*e.g.*, a chamber bowl 18), that improve the safety of synthesizers. The ventilation systems of the present invention may be applied to any type of synthesizer, and preferably, to open type synthesizers. These systems are particularly useful for improving the function and safety of certain commercially available synthesizers, such as the ABI 3900 Synthesizer.

During normal operations and without any malfunction, fumes are nonetheless are emitted from the chamber bowl of the 3900 machine when the synthesizer is opened for access by an instrument operator (*e.g.*, when the top cover or lid enclosure is opened to retrieve columns after synthesis is completed). These emissions can be significant. In some instances, instruments such as the 3900 may be installed inside chemical fume hoods to collect such emissions from normal operations. However, placing machines in chemical fume hoods is not practical for a number of reasons. For example, the presence of a large instrument within a chemical fume hood limits the use of the hood for other purposes. Removal of the instrument when the hood is needed for another purpose is impractical, since many synthesizers are

physically connected to external reagent reservoirs, gas tanks or other supply sources, making frequent removal and reinstallation prohibitively complex. Another problem with using chemical fume hoods to contain and remove emissions is that, using this approach, the number of synthesizers that can be used at one time is limited by the amount of hood space available. This prevents the use of many synthesizers in parallel, *e.g.*, in an array of synthesizers, and therefore limits high-throughput synthesis capability. What is needed are systems to properly vent synthesizers, such as the 3900, that do not require placing the machines in chemical fume hoods.

The present invention provides systems for collecting emissions from synthesizers without the use of a separate fume hood. The present invention comprises a synthesizer having an integrated ventilation system to contain and remove vapor emissions. By way of example, the integrated ventilation system of the present invention is described as applied to the components and features of open synthesizers like the Applied Biosystems 3900 instrument. However, this configuration is used only as an example, and the integrated ventilation systems are not intended to be limited to the 3900 instrument or to any particular synthesizer. One aspect of the invention is to collect and remove vapors when the instrument is open, *e.g.*, for access by the operator to the reaction chamber (Figures, 48C, and 49A-C). In one embodiment of the present invention, the integrated ventilation system comprises a ventilated workspace. Embodiments of an integrated ventilation system comprising a ventilated workspace as applied to the 3900 instrument are shown in Figures 48A-C, 49A-C and 50 A-B. Another embodiment is diagrammed in Figures 51 A and B.

In some embodiments, a ventilation opening is provided through an opening in the top. For example, referring to Figure 48A, in certain embodiments, some embodiments of synthesizers of the present invention comprise a top enclosure (*e.g.* 97) that forms a primarily enclosed space 104 over a top cover (*e.g.*, 30, not shown in this figure). In preferred embodiments, the top enclosure has four sides (*e.g.*, 98, two of which are shown in Figure 48A), and a top panel (*e.g.*, 99) that form a primarily enclosed space 104 above the top cover (*e.g.*, 30) containing a plurality of valves (*e.g.*, 10, not shown in this figure) and a plurality of dispense lines (*e.g.*, 6, not shown in this figure). In certain embodiments, the top panel (*e.g.*, 99) contains an outer window (*e.g.*, 101). In some preferred embodiments, the outer window contains a ventilation opening (*e.g.*, 105).

As used herein, the combination of a top enclosure (*e.g.*, 97) and top cover (*e.g.*, 30) is referred to collectively as the "lid enclosure" (*e.g.*, 102). In preferred embodiments, the "lid enclosure" has six sides, with the top cover (*e.g.*, 30) serving as the "bottom", the top panel serving as the surface opposite the top cover, and the four side walls being the top enclosure sides (*e.g.*, 98). In certain embodiments, the lid enclosure has a ventilation opening (*e.g.*, 105) with a ventilation tube (*e.g.*, 103) attached thereto (See, Figure 48B). In preferred embodiments, the ventilation tube is connected to a ventilation opening in an outer window 101.

In other embodiments, the synthesizer base (*e.g.*, 2) comprises a primarily enclosed space 104. In certain embodiments, a base (*e.g.*, 2) of a synthesizer comprises a ventilation opening (*e.g.*, 105) with a ventilation tube (*e.g.*, 103) attached thereto (See, *e.g.*, Figures 51A and 51B).

The ventilation openings in the lid enclosure or the base may be in any suitable position. For example, the ventilation opening in the lid enclosure may be in the top panel (e.g. in the center, toward the back of the machine, or in one of the corners). The ventilation opening may also be located in a top enclosure side. For example, the ventilation opening may be in the enclosure side at the back of the machine, or on one of the sides (*e.g.*, configured such that the lid enclosure may still be moved upward and downward while attached to a ventilation tube). A ventilation opening in a base may be, for example, on the front, the sides or on the back (*e.g.*, configured such that the lid enclosure may still be moved upward and downward without interference by the ventilation tube). In preferred embodiments, the ventilation opening is positioned toward the rear (*e.g.*, on a side or in the back) to allow the ventilation tubing to be directed away from an instrument operator. In particularly preferred embodiments, the ventilation opening is on the back of the base, *e.g.*, as shown in Figures 51A and 51B.

In some embodiments, the ventilation is located in a position such that air traveling through the primarily enclosed space (*e.g.*, 104) make greater or less contact with particular synthesizer components located inside the lid enclosure (e.g. valves, solenoids, dispense lines, etc.). The lid enclosures of the present invention may also have a plurality of ventilation openings. This may be desirable in order to control or direct air flow through the primarily enclosed space (*e.g.*, to minimize or to maximize air contact with particular synthesizer components inside the lid enclosure).

As shown in Figure 48C, in certain embodiments, the lid enclosure is hinged so that is may be moved upward and downward (*e.g.*, allowing access to the chamber bowl or other reaction chamber by a user). In some embodiments, the primarily enclosed space of the lid enclosure (e.g. 104, not shown in this figure) is open to the ambient environment through a ventilation slot (e.g. 100) in the top cover or the top enclosure (e.g. in top enclosure side wall towards the back of the machine).

In certain embodiments of the present invention, a lid enclosure is present on a commercially available machine (*e.g.*, ABI 3900), and the lid enclosure is modified as described herein (*e.g.*, a ventilation opening is made in the lid enclosure) An opening near the hinge for wiring serves as a ventilation slot on the 3900. In other embodiments, the lid enclosure must be added to synthesizer. For example, a synthesizer that simply has a top cover (*e.g.*, 30), may have a top enclosure (*e.g.*, 97) added thereto. This may be done by attaching a top enclosure that has bottom flanges (opposite the top panel) that fit around the top cover, and provide a point of attachment (*e.g.*, bolts, screws, adhesives, etc.). In other embodiments, the lid enclosure is fabricated as a separate component, then installed onto a synthesizer. For example, the components making up the lid enclosure (top enclosure and top cover) may be formed from a single mold, or two molds, etc. In this regard, features of the present invention may be built into the lid enclosure, such as the ventilation opening, ventilation slot, and certain hood components (described below).

In some embodiments, *e.g.*, as diagrammed in Figures 48A-C, the lid enclosure (*e.g.*, 102) comprises, or is modified to comprise at least one ventilation opening (*e.g.*, 105). One or more ventilation openings may be used. In preferred embodiments, a ventilation opening is placed in the center of the top panel so as to avoid blocking the operator's view of internal components, such as the synthesis columns, during operation. In preferred embodiments, the lid enclosure comprises windows constructed of transparent or translucent material, such as plexiglass.

In preferred embodiments, the lid enclosures of the present invention comprise a top panel directly opposite a top cover, and side walls between these two components The primarily enclosed space between the top panel and top cover is, in some embodiments, open to the ambient environment through a ventilation slot near the lid enclosure hinge (*e.g.*, 106). In

certain embodiments, the lid enclosure of the present invention comprises an inner window and an outer window (e.g. an outer window in the top panel, and an inner window in the top cover). The outer window of the instrument allows visual inspection of operations and components within the lid and within the chamber bowl 18 of the base 2. The inner window seals the chamber bowl 18 by pressing against the chamber gasket when the lid enclosure is closed. Reagent supply tubing passes through the inner window, but the window is sealed around each tube so that the chamber will maintain appropriate pressure during operation. In the embodiment shown in Figure 48B, the ventilation opening provides an aperture is the outer window.

In preferred embodiments, the ventilation opening (*e.g.*, 105) is attached to a ventilation tube (*e.g.*, 103), that in turn may be attached to an exhaust system. In some embodiments, a synthesizer is attached to an individual exhaust system. In other embodiments, multiple synthesizers are attached to a centralized exhaust system (e.g. centralized venting or vacuum system). In a preferred configuration, access to the exhaust system is toward the rear of the instrument, to minimize or prevent interference by the ventilation tubing with operator access to the chamber bowl, and to conduct the fumes away from instrument operators. The centralized exhaust may be a constant vacuum or a periodically actuated vacuum. In particular embodiments, raising the top cover or lid enclosure of a synthesizer triggers the vacuum system. In certain embodiments, reagent bottles on the sides of a synthesizer may also be vented through ventilation ports employing the same ventilation system employed by the ventilation tube attached to the top panel.

Another aspect of the present invention is to provide a ventilated workspace (*e.g.*, around the chamber bowl) having a negative air pressure relative to the surrounding air pressure, such that the flow of air goes from the surrounding room into the ventilated workspace, and not in the reverse, during operation of the ventilation system (*e.g.*, as shown in Figure 50B and 50B). The ventilated workspace is designed to allow the instrument operator to reach into the space (*e.g.*, to remove the synthesis columns) without turning off the ventilation system. One embodiment of a ventilated workspace is shown in Figure 49A, wherein the ventilated workspace is created by providing side panels (*e.g.*, 107). Two variations of another embodiment are shown in Figures 49B and 49C. In this embodiment, the ventilated workspace is created by providing side panels (*e.g.*, 107) between the body of the synthesizer and the lid enclosure, and a front panel (*e.g.*,

108). In certain embodiments, the ventilated workspace is created by including only side panels. In other embodiments, the ventilated workspace is created by only including a front panel. In preferred embodiments, side and front panels are used together (*e.g.*, as in Figures 49B and 49C) to create a ventilated workspace. In some embodiments, side and front panels are provided as separate components. In other embodiments, a single component comprising both side panels and a front panel is provided.

The size of the ventilated workspace can be altered by the placement of the panels, *e.g.*, the side panels (107) shown in Figures 49 A-C. In some embodiments, panels are positioned to maximize the size of the enclosed ventilated workspace (*e.g.*, as in Figure 49B). In other embodiments, the panels are positioned to provide a smaller ventilated workspace (*e.g.*, as with the side panels in Figure 49C). In some preferred embodiments, the side panels are positioned as close to the top chamber gasket (*e.g.*, 31) as they can be without disturbing the seal between the top chamber gasket and the top cover 30. In certain embodiments, the front and/or side panels are used with a synthesizer only having a top cover (not a full lid enclosure).

The side panels can be made of a number of different materials. In some embodiments, the materials used for the side panels are opaque. In other embodiments, the side panels are translucent or clear (*e.g.*, to permit surrounding light into the ventilated workspace). In certain embodiments, the side panels are constructed from flexible polymeric material (*e.g.*, sheeting), such as polyethylene or polypropylene. In some embodiments, the polymeric material has an average thickness of about 2 to 8 mils. In preferred embodiments, the polymeric material has an average thickness of about 2 to 4 mils. In some embodiments, the panels are collapsible (*i.e.*, can collapse or fold down upon themselves as the lid enclosure or top cover, is lowered). In some embodiments, panels are accordion-style or fan-fold style barriers that fold down upon themselves when the top cover or lid enclosure is lowered. In preferred embodiments, when the panels are collapsed, they have a total thickness that is less than the height of the O-ring or gasket (*e.g.*, top chamber seal 31) on the interior of the synthesizer (*e.g.*, so that there is no interference with the sealing of the O-ring).

In other embodiments, the side panels are constructed of rigid material. In some embodiments, rigid side panels are configured to fit into recesses in the body of the synthesizer when the top cover or lid enclosure is closed. In other embodiments, rigid side panels are

configured to fit around the outside of the base of the synthesizer when the top cover or lid enclosure is closed. In some embodiments, rigid side panels are constructed from opaque materials (*e.g.*, steel, aluminum, opaque plastic). In other embodiments, rigid side panels are constructed from translucent or transparent material, such as plexiglass. Generally, the side panels are connected to the top cover, so when the top cover or lid enclosure is raised, the side panels slide up to form sides for the ventilated workspace.

In certain embodiments, a front panel (*e.g.*, 108) is attached to the lid enclosure. For example, the front panel may attach to the top cover (*e.g.*, Figure 49B), or the front panel may attach to one of sides of the lid enclosure (*e.g.*, Figure 49C). The front panel may drape over the front of the synthesizer when the lid enclosure is closed (See, *e.g.*, Figures 48B and 49C). Alternatively, the front panel may fit into a recessed slot in the synthesizer base, or fold up upon itself as the lid enclosure is lowered into the closed position.

Attachment of the panels provided for the purpose of enclosing the ventilated workspace is not limited to any particular means. For example, in a simple configuration, panels are attached by use of strips of VELCRO fastener (*e.g.*, adhesive backed strips), for easy mounting and removal. For a sturdier attachment, the panels may be attached using fasteners, including but not limited to screws, bolts, welds, and snaps, or may be attached with removable or permanent adhesives. The presence of the panels reduces the size of the opening through which ambient air can enter the ventilated workspace, and also reduces the size of the opening from which air and vapors in the chamber bowl can escape. When the ventilation system is turned on (e.g., when the connected ventilation tube is drawing air from the ventilation opening, the airflow through the reduced opening prevents or reduces any flow (e.g. outward flow) of gaseous emissions. When the ventilation system is actuated, ambient air and reagent vapors are drawn across the chamber bowl (*e.g.*, 18) and into the ventilation slot (*e.g.*, 100), as diagrammed in Figures 50B and 51B. The air and vapors then move through the primarily enclosed space (*e.g.*, 104) and exit through the ventilation opening (*e.g.*, 105) into the ventilation tube (*e.g.*, 103). In some embodiments, the air flow rate at the opening of the ventilated workspace (*e.g.*, in the embodiments shown in Figures 49B and 49C, where the surrounding air is drawn into the ventilated workspace below the front panel and between the side panels) is from about 20 to

about 100 feet per minute, face velocity. In some preferred embodiments, the flow rate at the opening is about 40 to 50 feet per minute, face velocity.

From the ventilation tube, the air and vapors may be vented, treated or collected. In certain embodiments, the vented air and vapors are routed to a central scrubber. The central scrubber may form part of an overall emission control system. The central system may also be used to adjust total airflow for the number of synthesizers that are open at the same time. In this regard, exhaust from the system is minimized so as to concentrate waste vapors.

In order to increase or decrease the speed at which air and vapors travels through the ventilation system of the present invention, the size of the ventilation slot may be adjusted (e.g. reducing the size of the ventilation slot increase the speed of the moving air and vapors). The airflow pattern made possible by the present invention allows synthesizers to be opened (e.g. to change columns, etc) without exposure of an operator to hazardous vapors (e.g. argon, solvent fumes, etc).

The integrated chamber ventilation system of the present invention may be adapted to many synthesizers of both 'open' and 'closed' design. On example of another synthesizer that can be modified to include the reaction enclosure ventilation system of the present invention is the POLYPLEX 96-channel, high-throughput oligonucleotide synthesizer from GeneMachines, San Carlos, CA, which comprises a synthesis case providing an enclosure for the synthesis block in which the reactions are performed. A similar instrument is described in WO 00/56445, published September 28, 2000, and in related U.S. Provisional Patent application 60/125262, filed March 19, 1999, each incorporated herein in their entireties. As described in WO 00/56445, the synthesis case has a loading station, drain station, and water-tolerant and water-sensitive reagent filling stations. The synthesis case has a cover, a first and a second side, a first and a second end, and a bottom side, which contacts the base. The load station comprises a sealable opening in the synthesis case through which a multiwell plate can be inserted. In application of the present invention, the synthesis case can be fitted with one or more ventilation openings similar to ventilation opening 105, for attachment to ventilation tubing (*e.g.*, 103). In some embodiments, a ventilation opening is in a side of the synthesis case opposite the side having the sealable opening. In preferred embodiments, a ventilation opening in the synthesis case is on the first or second end. In particularly preferred embodiments, the ventilation system

is actuated when the sealable opening is opened, *e.g.*, for insertion or removal of a multiwell plate.

The present invention also contemplates robotic means (e.g. conveyor belt, robots, etc) for linking the synthesizers to other components of the production process. For example, Figure 52 illustrates a synthesizer 1, a robotic means 92, a cleave and deprotect component 93 and a purification component 94 operably linked together.

The present invention provides synthesizer arrays (*e.g.*, groups of synthesizers). In some embodiments, the synthesizers are arranged in banks. For example, a given bank of synthesizers may be used to produce one set of oligonucleotides. The present invention is not limited to any one synthesizer. Indeed, a variety of synthesizers are contemplated, including, but not limited to the synthesizers of the present invention, MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), OligoPilot (Amersham Pharmacia,), and the 3900 and 3948 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA). In some embodiments, synthesizers are modified or are wholly fabricated to meet physical or performance specifications particularly preferred for use in the synthesis component of the present invention. In some embodiments, two or more different DNA synthesizers are combined in one bank in order to optimize the quantities of different oligonucleotides needed. This allows for the rapid synthesis (*e.g.*, in less than 4 hours) of an entire set of oligonucleotides (all the oligonucleotide components needed for a particular assay, *e.g.*, for detection of one SNP using an INVADER assay [Third Wave Technologies, Madison, WI]).

In some embodiments the DNA synthesizer component includes at least 100 synthesizers. In other embodiments, the DNA synthesizer component includes at least 200 synthesizers. In still other embodiments, the DNA synthesizer component includes at least 250 synthesizers. In some embodiments, the DNA synthesizers are run 24 hours a day.

Synthesizer Example 1

The Northwest Engineering 48-Column Oligonucleotide Synthesizer

The Northwest Engineering 48-Column Oligonucleotide Synthesizer (NEI-48, Northwest Engineering, Inc., Alameda, CA) is an "open system" synthesizer in that the dispensing tubes for the delivery of reagents are not affixed to each synthesis vial or column for the entire term of the

synthesis process. Instead, movement of a round cartridge containing the columns allows each dispensing tube to serve multiple columns. In addition, when a synthesis column is positioned to receive reagent, the dispenser is not even temporarily affixed to the vial with a sealed coupling. The reagent dispensed to the vial has open contact with the surrounding environment of the chamber. The chamber containing the synthesis vials is isolated from the ambient environment by a top plate. The general design and operation of the NEI instrument is described in WO 99/656602.

The NEI-48 synthesizer includes external mounting points for various reagent bottles, such as the phosphoramidite monomers used to form the polymer chain, and the oxidizers, capping reagents and deblocking reagents used in the reaction steps. TEFLON tubing feeds liquid from each reagent bottle to its assigned valve on the top of the machine. The feeding is done under pressure from an argon gas source.

The operations of the machine are controlled using a computer. The computer is fitted with a motion control card connected via cabling to a motor controller in the synthesizer; in addition, the computer is connected to the synthesizer via an RS-232C cable. The provided software allows the user to monitor and control the machine's synthesis operations.

The machine also requires connection to a source of argon gas, to be delivered at a pressure between 15 and 60 psi, inclusive, and a source of compressed air or nitrogen, to be delivered at a pressure between 60 and 120 psi, inclusive.

Synthesis in the NEI-48 occurs within synthesizer columns that are arranged in the cartridge.

Operations of the NEI-48 in accordance with the manufacturer's instructions produced undesirable emissions and leakage resulting in potential synthesis and instrument failure. The following section details two of the sources of these emissions, and details one or more aspects of the present invention applied to solve each problem, to thereby improve the performance of this machine.

A. Column overflow due to inadequate argon pressure

Undesirable emissions and exposure are increased when columns overflow, causing the hazardous reagents used during synthesis to collect in the chamber bowl. A number of types of

malfunction in the machine can leads to incomplete drainage or purge of the columns, and each will eventually lead to column overflow as the instrument proceeds through its subsequent dispensing steps.

The flow of reagent and waste from the synthesis columns is controlled by a differential in the pressure of argon between the top and bottom openings of the column. When the pressure of argon on the top opening is not sufficiently high, the column will not drain or be purged completely, *i.e.*, fluid that should be drained will remain in the column. This improper purging not only reduces the efficiency of the synthesis chemistry, it also leads to column overflow. Therefore, failure of either initial pressurization of the chamber, or leakage of argon from any coupling (in an amount great enough to reduce either the overall pressure of the system or the pressure differential across the synthesis column) may lead to undesirable emissions and exposure. One aspect of the present invention is to prevent column overflow by reducing leakage of argon at a variety of points in the system.

The NEI-48 demonstrated a variety of failures as a result of argon leakage from or within the instrument. To address this problem, the drain plate gasket 43 of the present invention was created and was fitted between the cartridge and drain plate. Addition of the gasket to this assembly, as diagramed in Figure 38, provided a pressure-tight seal, thereby containing the argon and allowing proper drainage of the columns at the purging step. The gasket of the present invention applied in this way improved the safety of the machine, and improved the efficiency of the synthesis reaction.

In another embodiment, a modified drain plate gasket was provided. The drain plate has securing holes 33, for attachment of the motor connector 22. The first gasket was of a design that avoided the areas of the motor connector 22 and the securing holes 33. A modified drain plate gasket was designed with guide holes 44 to fit closely around each securing hole 33, such that the holes served to place the gasket in a specific position between the cartridge and the drain plate (Figure 38). In an alternative embodiment, the drain plate 19 and the cartridge 3 may be provided with other alignment features, such as pin fittings and corresponding pin receiving holes (not shown) to facilitate alignment of these parts during assembly (*e.g.*, after cleaning). A modified drain plate gasket for use with these parts may be provided with pin guide holes (not shown). Use of either the securing holes 33, or pins fittings to align the gasket makes the gasket

easier to position during assembly, ensuring proper operation of the gasket and improving ease of any maintenance that requires disassembly of these parts.

B. Emissions from reagent bottles

During normal operations and without any malfunction, fumes can nonetheless be emitted by the reagent bottles attached to the machine. These emissions can be increased by poor fit or incorrect seals around bottle caps. For example, the reagent bottles for the NEI-48 are affixed to the machine by clamps that apply pressure to the outside of the bottle caps. The clamps can distort the caps, increasing leakage and gaseous emissions.

One aspect of the present invention is to provide a means of collecting emissions from reagent bottles. For improving the NEI-48, a reagent stand comprising a ventilation tube was constructed. The stand holds the reagent bottles, thereby eliminating the need for the cap-distorting clamps, and consequently reducing emissions from the bottles; the ventilation tube removes any remaining emitted gases. This reagent dispensing station improves the safety of the machine in normal operation. The reagent dispensing station of the present invention is not limited to a configuration comprising a stand. It is envisioned that a station comprising a ventilation system may also be used with one or more bottles held in clamps. In preferred embodiments, at least one aspect of the reagent container system, *e.g.*, the clamp, the cap, or the bottle, is modified such that clamping the reagent bottle does not compromise the containment function of the cap, or of any other aspect of the reagent container system.

Synthesizer Example 2

The Applied Biosystems 3900 Oligonucleotide Synthesizer

The Applied Biosystems 3900 Oligonucleotide Synthesizer (Applied Biosystems, Foster City, CA) is similar in design and function to the NEI-48, described above. The 3900 is an "open system" synthesizer utilizing a round cartridge containing the columns. The receiving holes of the cartridge are essentially cylindrical, and, as with the NEI-48, proper function of the instrument relies on an airtight seal between the columns and cartridge.

The 3900 synthesizer includes recessed areas for the external mounting of reagent bottles. When mounted on the instrument, the reagent bottles do not protrude beyond the outside edges

of the instrument; they are completely recessed, (as, *e.g.*, the reagent reservoirs 72 are recessed in base 2, diagrammed in Figure 47A). As with the NEI-48, the reagent feeding is done under pressure from an argon gas source.

The performance of the 3900 synthesizer is improved using the modifications provided by the present invention. Two specific improvements are described below. These particular improvements are described by way of example; improvements to the ABI 3900 synthesizer, or any synthesizer, are not limited to the improvements described herein below.

A. Column overflow due to inadequate argon pressure

As described above for the NEI-48, the proper purging of the synthesis columns at each cycle relies on the maintenance of a differential in argon pressure between the top and bottom openings of the columns. Improper or incomplete purging reduces the efficiency of the synthesis and increases the risk of column overflow. Proper purging in the 3900, like other open systems, depends in part upon the formation of an airtight seal between receiving holes in the cartridge and exterior surfaces of the synthesis columns. The presence of irregularities in the column shape or surface can prevent the formation of an airtight seal, allowing argon to leak around the column exterior, thereby disrupting the pressure differential required to properly purge the columns at each cycle. The need to discard columns having even minor imperfections adds expense to the use of the instrument. If undetected, a faulty seal can lead to poor synthesis and column overflow, as described above.

As discussed above, in some embodiments, the present invention provides improved synthesizers having reliable seals between the cartridge and the synthesis columns. The present invention provides a number of embodiments of synthesizers having such seals. For example, as described above, a synthesizer may be improved by the addition of a resilient seal, such as an O-ring, in the receiving hole of each cartridge.

To make this improvement, the 3900 is fitted with such O-rings for safer, more reliable and more efficient performance. Examples of several means of creating an improved seal between the outer surface of a column 61 and a receiving hole 11 are diagrammed in Figures 46A-46C. While any of the embodiments of seals disclosed herein may be applied to the 3900 instrument, in a preferred embodiment, the 3900 is improved by the use of an embodiment

similar to that diagrammed in Figure 46B, wherein a groove 70 creates a groove lip 71, to accommodate and hold an O-ring 67, thus providing a seal between cartridge 3 and the exterior surface 61 of the synthesis column 12. In a particularly preferred embodiment, the receiving hole 11 is enlarged in diameter to facilitate insertion and removal of an O-ring 67, *e.g.*, for easy cleaning or replacement. A groove is machined into the interior of each receiving hole in a 3900 cartridge, and appropriate O-ring seals are placed in the grooves. As noted above, the O-ring could be of any suitable material. Thus modified, the cartridge of the 3900 has a greatly improved ability to accommodate imperfections in the exteriors of synthesis columns, and this improvement results in safer, and more efficient and reliable operation of the instrument, with fewer costs associated with chemical spill clean-up, instrument down-time, and the disposal of unusable synthesis columns.

B. Emissions from reagent bottles

During normal operations and without any malfunction, fumes are nonetheless emitted by the reagent bottles attached to the 3900 machine. These emissions can be significant, even though gaskets are provided for use in conjunction with the bottle caps.

As described above, the present invention provides a means of collecting emissions from reagent bottles. On the 3900, the reagent bottles are attached in recessed areas on the exterior in the base of the instrument (*e.g.*, the reagent reservoirs 72 attached to the recessed areas in the base 2, as illustrated in Figure 47A). The emissions from this instrument are reduced by modification to provide the enclosed reagent dispensing station of the present invention. In modification of the 3900, the recessed areas are provided with panels to enclose the space, reducing the release of hazardous vapors.

Reagent bottles or reservoirs need to be accessible for changing or filling, due, *e.g.*, to consumption of reagents during synthesis operations. In making the modification to the 3900, the panels added to the instrument are moveable, to provide access to the reagent bottles within the enclosed space. In a simple configuration, panels provided for the purpose of enclosing the space are attached by use of strips of VELCRO fastener (*e.g.*, adhesive backed strips), for easy mounting and removal. For a sturdier attachment, the panels may be attached using hard, removable fasteners, such as screws or bolts. In a particularly preferred configuration, the panels

are mounted in tracks, brackets or other suitable fittings that allow them to be moved or removed by sliding.

To monitor reagent bottles (*e.g.*, to determine when changing or filling is needed), it is preferred that the reagent reservoirs be accessible for visual inspection. In making the addition of panels to the 3900, the panels are constructed such that the reagent bottles can be visually inspected without opening the enclosure. The panels provided are constructed of transparent material. While glass may be used, in preferred embodiments, for both safety and ease of handling a plastic is used with sufficient transparency to allow visual inspection of reagent bottles, and with sufficient resistance to the chemicals used in synthesis to avoid rapid or immediate decay or fogging, (as is often associated with exposure of plastics to vapors of solvents to which they are not resistant), when used in this application. Selection of plastics for appropriate chemical resistance is well known in the art, and tables of chemical compatibility are generally readily available from manufacturers.

The panels are provided with a ventilation port (*e.g.*, ventilation port 74, as diagrammed in Figure 47B), for the removal vapors and fumes emitted by the reagent bottles. Such a ventilation port serves as an attachment point for a ventilation tube to conduct fumes away from the instrument, *e.g.*, into an exhaust system. Since the vapors from DNA synthesis reagents tend to be heavier than air, the ventilation port is placed near the bottom of the enclosure. Placement of the ventilation port toward the rear is convenient for attachment to a larger exhaust system, minimizes or prevents interference by the ventilation tubing with operator access to other parts of the instrument, and conducts the fumes away from instrument operators.

To maximize efficacy of the ventilation system, an air inlet into the enclosure is provided. In applying the panels to the 3900, a clearance between the attached panels and the body of the instrument (*e.g.*, the clearance 75 between the panel 73 and the base 2 diagrammed in Figure 47B) provides the air inlet. The panel is positioned such that the principal air inlet is a clearance between the front edge of the panel (*i.e.*, the edge closest to the front of the instrument) and the instrument base. Positioning of the inlet toward the front of the instrument, or on the opposite side of an enclosure from a ventilation port, maximizes the flow of air through the enclosure, providing the most efficient removal of vapors. The inward flow of air minimizes the possible

escape of hazardous vapors toward instrument operators. Thus modified, the 3900 instrument is improved with respect to its emissions of hazardous vapors.

C. Emissions from the chamber bowl

During normal operations and without any malfunction, fumes are nonetheless emitted when the chamber bowl of the ABI 3900 is opened for access by the instrument operator (*e.g.*, when the lid is opened to retrieve columns after synthesis is completed). These emissions can be significant. The present invention provides a means of collecting emissions from the 3900 without the use of a separate fume hood. The present invention comprises a synthesizer having an integrated ventilation system to contain and remove vapor emissions. One aspect of the invention is to collect and remove vapors when the instrument is open. Embodiments of integrated ventilation systems as applied to the 3900 instrument are shown in Figures 48-51.

As shown in Figure 48A, in one embodiment, the lid enclosure 102 is modified to comprise a ventilation opening 105. The lid enclosure of the 3900 comprises an outer window 101. In preferred embodiments, a ventilation opening is placed in the center of the outer window 101 of the lid enclosure 105, so as to avoid blocking the operator's view of internal components, such as the synthesis columns, during operation.

As shown in the diagram of Figure 50, the lid enclosure of the 3900 instrument comprises an outer window 101 and an inner window 25. The space between the windows is open to the ambient environment through a ventilation slot 100 near the lid enclosure hinge 106. The outer window in an unmodified instrument allows visual inspection of operations and components within the lid enclosure and within the chamber bowl 18 of the base 2. Reagent supply tubing passes through the inner window, but the window is sealed around each tube so that the chamber will maintain appropriate pressure during operation. In the embodiment shown in Figures 48, 49 and 50, the ventilation opening provides an aperture in the outer window.

In another embodiment, one or more ventilation openings may be provided in the base (*e.g.*, 2) of the synthesizer, as diagrammed in Figure 51. In other embodiments, a synthesizer may comprise ventilation openings in both a lid enclosure and a base.

Each ventilation opening is attached to ventilation tubing (*e.g.*, 103) for attachment to an exhaust system. In some embodiments, a synthesizer is attached to an individual exhaust system.

In other embodiments, multiple synthesizers are attached to a centralized exhaust system. In a preferred configuration, the access to the exhaust system is toward the rear of the instrument, to minimize or prevent interference by the ventilation tubing with operator access to the chamber bowl, and to conduct the fumes away from instrument operators.

Another aspect of the present invention is to provide a ventilated workspace around the chamber bowl having a negative air pressure relative to the surrounding air pressure, such that the flow of air goes from the surrounding room into the ventilated workspace, and not in the reverse, during operation of the ventilation system. The ventilated workspace is designed to allow the instrument operator to reach into the space (*e.g.*, to remove the synthesis columns) without turning off the ventilation system. Embodiments of a ventilated workspace are shown in Figure 49 A-C. As shown in this embodiment, the ventilated workspace is created by providing side panels between the body of the synthesizer and the lid enclosure, and a front panel. The presence of the panels reduces the size of the opening through which ambient air can enter the ventilated workspace. When the ventilation system is turned on (*i.e.*, when the connected ventilation tube is drawing air from the ventilation opening, the airflow in through the reduced opening prevents or reduces any outward flow of gaseous emissions.

B. Closed System Synthesizers

In preferred embodiments, the present invention provides closed-system solid phase synthesizers that are suitable for use in large-scale polymer production facilities. Each synthesizer is itself capable of producing large volumes of polymers. Furthermore, the present invention provides systems for integrating multiple synthesizers into a production facility, to further increase production capabilities.

Currently available nucleic acid synthesizers have limited synthesis capacity. For example, the 3900 DNA Synthesizer (Applied Biosystem, Foster City, CA) is one of the most capable synthesizers and produces fewer than 100 40-mer oligonucleotides in a typical day production run. Additional synthesizers are described in U.S. Pat. Nos. 5,744,102, 4,598,049, 5,202,418, 5,338,831, 5,342,585, 6,045,755, and 6,121,054, and PCT publication WO 01/41918, herein incorporated by reference in their entireties.

The synthesizers of the present invention dramatically increase capacity, with some embodiments allowing over 2000 40-mer oligonucleotides to be produced per day (e.g., during a 16 hour production day) at a 1 μM scale. These capacities are achieved through the use of multi-chamber reaction supports that allow parallel synthesis of polymers within each chamber. For example, three or more chambers (e.g., comprising synthesis columns), preferably 96 or more chambers are provided on a reaction support, permitting a plurality of different oligonucleotides to be simultaneously produced. Each reaction chamber is associated with its own reagent dispenser such that reagents are delivered to each chamber substantially simultaneously rather than delivery reagents in sequence. In preferred embodiments, the synthesizer is a closed system during operation (i.e., reagent delivery to the chambers and waste removal from the chambers occurs in a continuous pathway that is isolated from the ambient environment). An example of a closed system is illustrated in Figure 53. In some preferred embodiments, the synthesizers have a minimum number of moving parts. In particular, the reaction support is immobile.

In some embodiments, the synthesizer provides additional polymer production capabilities. For example, in some embodiments, the synthesizer is configured to conduct cleavage and deprotection of synthesized oligonucleotide. In preferred embodiments, the same reaction support is used for both synthesis and cleavage and deprotection. In other preferred embodiments, the same reagent dispensers are used for both synthesis and cleavage and deprotection. In still other preferred embodiments, the reaction support does not move during both the synthesis and cleavage and deprotection processes (i.e., synthesis and cleavage and deprotection occur at the same location). In some embodiments, the synthesizer also provides an integrated purification component (e.g., using the same reaction support and/or reagent dispensers with or without movement of the reaction support). Any other production components described herein may also be integrated with the synthesizer.

Preferred features of the synthesizers of the present invention include: single day synthesis capacities of 2000 oligonucleotides, based on an average 40-mer at 1 μM scale with 16 hours staffing; production scale capabilities of 40, 100, 1000, and 4000 nM, with larger scales supported by control elements; compatibility with commercially available nucleic acid synthesis columns (e.g., columns designed for use with EXPEDITE nucleic acid synthesizers [Applied Biosystems, Foster City, CA], 3900 High-Throughput Columns for use with the 3900 DNA

Synthesizer [Applied Biosystems], DNA synthesis columns from Biosearch Technologies, Novato, CA); mechanical and/or data interface capability with other production components (see Section II, below); individual oligonucleotide tracking (e.g., during synthesis and throughout an entire production process); compatibility with standard nucleic acid synthesis chemistry with provisions for optimization of reaction conditions; detectors for monitoring trityl or other components or reagents; compatibility with standard multi-chamber formats (e.g., 96-well plate, 384-well plate formats); interface with databases to input and track information including, but not limited to oligonucleotide sequence, completion, data, time, and channel; and integration with a control system to allow multiple synthesizers to have a common control center.

Reagent delivery to the synthesizer is achieved using a novel fluidics system. In preferred embodiments, all fluid transfers are desired to be closed system; that is, a closed fluid circuit exists from source to waste at any time reagents are being transferred. In general, the supply circuit remains coupled to the synthesis columns that are supported by the reaction support for all operations except, in some embodiments, during nucleic acid coupling reactions. Given the reaction time required for the coupling reactions (approximately 30 seconds), in some embodiments, the circuit to a particular column or columns is disconnected to allow fluid transfer mechanisms to be used on other columns. While the fluid transfer is re-routed, the columns undergoing the coupling reaction need not be exposed to the ambient environment (i.e., a sealed delivery path may be maintained).

In preferred embodiments, the target fluid transfer system is a pressurized supply with dispense control valves. Reagents flow to the reaction chambers upon opening of the control valves, driven by a pressure differential.

In some preferred embodiments, the reaction support contains waste channels configured to receive waste from the reaction chambers. In some embodiments, each channel is configured with its own waste channel (See e.g., Figure 53). The waste channels preferably feed into a single waste disposal line. In some embodiments, the waste system is gravity driven. In other embodiments, a valve-controlled vacuum is used to eliminate waste. In some preferred embodiments, waste lines are fitted with a trityl monitoring device. In preferred embodiments, the waste line is fitted with a qualitative trityl monitoring device. For example, colorimetric analysis of effluent using a CCD camera or a similar device provides a yes/no answer on a

particular detritylation level. Qualitative detection of detritylation can generally be performed with less expensive equipment than is generally required by more precise quantitation, and yet generally provides sufficient monitoring for detritylation failure. Valves used to control reagent delivery and/or waste removal may be under automated control.

In preferred embodiments, a plurality of reagent dispensers are provided, wherein a reagent dispenser is provided for each reaction chamber. In such embodiments, the reagent dispensers provide each of the reagents necessary to support a synthesis reaction within the reaction chamber. For nucleic acid synthesis, this includes, for example, delivery of acetonitrile, phosphoramidite corresponding to each of the bases, argon gas, oxidizer, activator (*e.g.*, tetrazole), deblocking solution and capping solution. Thus, in some embodiments, the reagent dispenser comprises a plurality of reagent delivery lines, each line providing a direct fluidic connection between the reagent dispenser and individual supply tanks for the different reagents (See *e.g.*, Figure 53).

An example of such a reagent dispenser (2) is shown in Figure 54 from both a side view (Figure 54A) and a cross-sectional bottom view (Figure 54B). The side view shows a single reagent delivery line (3) penetrating a top surface (4) of the reagent dispenser (2). In this embodiment, a retention ring (5) is used to support the reagent delivery line (3). The reagent delivery line (3) ends at a reagent reservoir (6) that is configured to receive reagents from each of the delivery lines. A seal (7) forms a contact between the delivery line (3) and the reagent reservoir (6). The center of the reagent reservoir (6) comprises a delivery aperture (8). The delivery aperture (8) is in fluidic contact with a delivery channel (9), with a seal (10) forming a contact between the delivery channel (9) and the delivery aperture (8). The delivery channel (9) passes through a bottom surface (11) of the reagent dispenser (2) and may positioned by a retention ring (12).

The cross-sectional bottom view shown in Figure 54B shows the presence of nine delivery lines (3) contained within the reagent dispenser (2). Each delivery line empties into the reagent reservoir (6), represented by the eight pronged star. Figure 55A shows one preferred embodiment of the reagent dispenser (2), wherein the outer surface of the delivery channel (9) contains first (13) and second (14) ring seals configured to form an airtight or substantially airtight seal with one or more points on the interior surface of a synthesis column (15) or other

reaction chamber (e.g., with reaction chambers present in a synthesizer or a cleavage and deprotection component; see, for example Figure 55B).

In preferred embodiments, common reagent tanks supply reagents to all of the reaction chambers. The reagents tanks may be contained within the synthesizer or may be external to the synthesizer. Where the tanks are provided with the synthesizer, they are preferably contained in a vented chamber to reduce the build-up of gaseous or liquid waste in and around the synthesizer. In some preferred embodiments, common reagent tanks supply reagents to a plurality of synthesizers. Examples of such delivery systems are provided, below. In yet other embodiments, some of the reagents are supplied externally and some of the reagents are supplied at or in the synthesizer (*e.g.*, amidites). In some embodiments, one or more of the reagents are processed, *e.g.*, under vacuum, to remove dissolved gasses.

In some preferred embodiments, the synthesizer comprises a means of delivering energy to the reaction chambers to, for example, increase nucleic acid coupling reaction speed and efficiency, allowing increased production capacity. In some embodiments, the delivery of energy comprises delivering heat to the reaction chambers. In addition to increasing production capacity, the use of heat allows the use of alternate synthesis chemistries and methods, *e.g.*, the phosphate triester method, which has the advantages of using more stable monomer reagents for synthesis, and of not using tetrazole or its derivatives as condensation catalysts. Heat may be provided by a number of means, including, but not limited to, resistance heaters, visible or infrared light, microwaves, Peltier devices, transfer from fluids or gasses (*e.g.*, via channels or a jacketed system). In some embodiments, heat generated by another component of a synthesis or production facility system (*e.g.*, during a waste neutralization step) is used to provide heat to reaction chambers. In other embodiments, heat is delivered through the use of one or more heated reagents. Delivery of heat to reaction chambers also comprises embodiments wherein heat is created within the reaction chamber, *e.g.*, by magnetic induction or microwave treatment. It is contemplated that heating may be accomplished through a combination of two or more different means.

In some embodiments, the delivery of heat provides substantially uniform heating to two or more reaction chambers. In some embodiments, heating is carried out at a temperature in a range of about 20 °C to about 60 °C. The present invention also provides methods for

determining an optimum temperature for a particular coupling chemistry. For example, multiple synthesizers are run side-by-side with each machine run at a different temperature. Coupling efficiencies are measured and the optimum temperature for one or more incubations times are determined. In other embodiments, different amounts of heat are delivered to different reaction chambers within a single synthesizer, such that different reaction chemistries or protocols can be run at the same time.

Delivery of heat to a closed system will alter the pressure within the system. It is contemplated that the closed system of the present invention will be configured to tolerate variations in the system pressure (*i.e.*, the pressure within the closed system) related to heating or other energy input to the system. In preferred embodiments, the system (*e.g.*, every component of the system and every junction or seal within the system) will be configured to withstand a range of pressures, *e.g.*, pressures ranging from 0 to at least 1 atm, or about 15 psi. It is contemplated that pressures may be varied between different points within the system. For example, in some embodiments, reagents and waste fluids are moved through the reaction chamber by use of a pressure differential between one end (*e.g.*, an input aperture) and the other (*e.g.*, a drain aperture) of the reaction chamber. In some embodiments, the system of the present invention is configured to use pressure differentials within a pressurized system (*e.g.*, wherein a system segment having lower pressure than another system segment nonetheless has higher pressure than the environment outside the closed system). In some embodiments, the prevention of backward flow of reagents through the system (*e.g.*, in the event of back pressure from a process step such as heating) is controlled by use of pressure. In other embodiments, valves are provided to assist in control of the direction of flow.

In other preferred embodiments, the synthesizer comprises a mixing component configured to mix reaction components, *e.g.*, to facilitate the penetration of reagents into the pores of the solid support. Mixing may be accomplished by a number of means. In some embodiments, mixing is accomplished by forced movement of the fluid through the matrix (*e.g.*, moving it back and forth or circulating it through the matrix using pressure and/or vacuum, or with a fluid oscillator). Mixing may also be accomplished by agitating the contents of the reaction chamber (*e.g.*, stirring, shaking, continuous or pulsed ultra or subsonic waves, See, Figures 42A-C and 43A and B). In some preferred embodiments, an agitator is used that avoids

the creation of standing waves in the reaction mixture. In some preferred embodiments, the agitator is configured to utilize a reaction vessel surface or reaction support surface (e.g., a surface of a synthesis column) to serve as resonant members to transfer energy into fluid within a reaction mixture. In some embodiments, the matrix is an active component of the mixing system. For example, in some embodiments, the matrix comprises paramagnetic particles that may be moved through the use of magnets to facilitate mixing. In some embodiments, the matrix is an active component of both mixing and heating systems (*e.g.*, paramagnetic particles may be agitated by magnetic control and heated by magnetic induction). It is contemplated that any of these mixing means may be used as the sole means of mixing, or that these mixing components may be used in combination, either simultaneously or in sequence. In preferred embodiments, the heating component and the mixing component are under automated control.

In preferred embodiments, a central control processor is used to automate one or more of the synthesis steps or synthesizer operations. The central control processor may also be configured to interact with one or more other components of a production facility (See below). In some embodiments, the central control processor regulates valves, controlling the timing, volume, a rate of reagent delivery to the reaction chambers. In preferred embodiments, all delivered reagents are controllable for volume within prescribed ranges at each step of the synthesis process within a protocol independent of other steps.

The present invention is not limited by the range of flow rate used for reagent delivery. However, in preferred embodiments, flow rates are 300-500 μL/sec for all reagents.

Table 1, below, provides an example of reagent delivery times (in seconds) and amounts (in microliters) for a single synthesis cycle. Conditions are provided for four different synthesis scales.

TABLE 1

Step	40 nM scale	200 nM scale	1 μM scale	4 μM scale	Time (sec)
add acetonitrile	50	150	250	1000	0.5
argon purge					1
add deblock	50	150	250	1000	0.5
argon purge					1

add deblock	50	150	250	1000	0.5
argon purge					1
add deblock	50	150	250	1000	0.5
argon purge					1
add deblock	50	150	250	1000	0.5
argon purge					1
add acetonitrile	50	150	250	1000	0.5
argon purge					1
add amidite and	15	30	75	300	30x4
tetrazole	20	45	115	460	
argon purge					1
add cap a	15	30	60	180	1
add cap b	15	30	60	180	
argon purge					1
add oxidizer	40	80	180	360	0.5
argon purge					1
add acetonitrile	100	200	250	1000	
argon purge					

In preferred embodiments, with the exception of the amidite coupling step, reaction or wash times are controlled by fluid application rate without additional dwell time prior to purging. This is in contrast to methods used with current commercial synthesizers (e.g., 3900 DNA Synthesizers).

A number of different configurations of the synthesizers of the present invention are provided below with exemplary capacities provided. The present invention is not limited to these specific configurations.

A. Pure batch, fully dedicated fluidics

Batch size is preferably 96 arrayed reaction chambers in a standard microtiter footprint. Synthesis columns could be either independently filled and inserted into a rack to form the array

or, preferably, molded in an arrayed format and filled as a batch. If the latter, then all columns should be of a similar type and synthesis operations are grouped accordingly. Column plates are loaded one at a time and replaced at the end of the synthesis process. In some embodiments, loading and unloading is manual--no transport mechanisms required. In other embodiments, loading and unloading is controlled robotically. Fluid connections from the system to the column tray is either established by the system (moving mechanism) or by the user en mass (fixed dispense). Application of reagents is accomplished by a fixed set of multifunctional reagent dispensers, each incorporating all required reagents: each column has a dedicated multiplexed supply line and no motion devices or fluid connection make/break cycles are required. This approach requires a large number of valves (approximately 1000) and is therefore preferably uses very compact, relatively inexpensive and relatively high reliability valves.

Estimated walk away time:	35 minutes
Optimal output per day:	approximately 2496 40-mers
Valve count:	1000
Mechanism level:	none
Size:	smallest

B. Pure Batch: non-dedicated fluidics

This system is similar to the pure batch system, but rather than dedicated fluidics for each channel, moving reagent dispense heads are provided. This reduces the valve count but adds mechanism. Also, output per day drops in some scale to the valve reduction. A system with approximately 200 valves would produce about 1056 oligonucleotides/2 shift day. Adding a parallel processing station to achieve 2112/day is an option. Walk away time goes up to approximately 80 minutes.

Estimated walkaway time:	1.3 hours
Optimal output per day:	approximately 2112 40-mers
Valve count:	400
Mechanisms level:	moderate

Size:	moderate

C. Modified Batch:

This system is similar in configuration to the non-dedicated fluidics batch system described above, but allows multiple plate positions with the system. Walkaway time improves linearly with the number of plates allowed, throughput and other comments are similar. At increasing levels of resident plates, parallel (400 valve system) with 4 plates resident for each parallel line would allow walk away time of 5 hours. In principle, 4 runs of 8 plates could be completed per day producing 3072 oligonucleotides. A 200-valve system configured similarly could produce 1536.

Estimated walkaway time:	5 hours
Optimal output per day:	approximately 1536 40-mers
Valve count:	200
Mechanism level:	moderate
Size:	moderate

D. Continuous Batch:

This system is similar to the above system with the addition of queues for feeding plates and accumulating completed plates. The system requires similar fluid handling but adds plate transport mechanisms. The waste system is more complicated due to plate movement. This system allows direct integration to downstream cleave and deprotect system and allows direct integration to synthesis column packing upstream. Throughput is slightly higher than the modified batch system.

Estimated walkaway time:	Limited only by onboard storage
Optimal output per day:	approximately 1536 40-mers
Valve count:	200
Mechanism level:	high
Size:	large

E. Continuous Parallel:

Rather than a 96-well format, the columns are prepared and presented in strips of 12 columns. The strips are fed through multiple parallel reagent delivery ports. This approach allows greater spacing between adjacent fluidic elements and allows processing of multiple different column types simultaneously. An additional benefit is the likelihood that a closer approach to the theoretical maximum throughput should be routinely achieved. In this embodiment, throughput per valve would be similar to continuous batch, but tubing of throughput is easier.

Estimated walkaway time:	limited only by onboard storage
Optimal output per day:	approximately 1536 40-mers
Valve count:	200
Mechanism level:	high
Size:	large

(All valve counts are approximate and assume 2 way valves: with multi-position valves, the counts drop accordingly. Also, some rejection may be possible by ganging operations less critically dependent on precise fluid delivery (washes etc). All throughputs assume a nominal cycle for 1 uM scale. Larger scale(s) would be significantly longer. Smaller scales would be essentially similar. Mixing longer and shorter oligonucleotides will drive throughputs to that presented by the longer oligonucleotides).

The synthesizers of the present invention also provide components to reduce or eliminate undesired emissions. A problem with currently available synthesizers is the emission of undesirable gaseous or liquid materials that pose health, environmental, and explosive hazards. Such emissions result from both the normal operation of the instrument and from instrument failures. Emissions that result from instrument failures cause a reduction or loss of synthesis efficiency and can provoke further failures and/or complete synthesizer failure. Correction of failures may require taking the synthesizer off-line for cleaning and repair. The present invention provides nucleic acid synthesizers with components that reduce or eliminate unwanted

emissions and that compensate for and facilitate the removal of unwanted emissions, to the extent that they occur at all. The present invention also provides waste handling systems to eliminate or reduce exposure of emissions to the users or the environment. Such systems find use with individual synthesizers, as well as in large-scale synthesis facilities comprising many synthesizers (*e.g.* arrays of synthesizers).

Whether a system used is open or closed, oligonucleotide synthesis involves the use of an array of hazardous materials, including but not limited to methylene chloride, pyridine, acetic anhydride, 2,6-lutidine, acetonitrile, tetrahydrofurane, and toluene. These reagents can have a variety of harmful effects on those who may be exposed to them. They can be mildly or extremely irritating or toxic upon short-term exposure; several are more severely toxic and/or carcinogenic with long-term exposure. Many can create a fire or explosion hazard if not properly contained. In addition, many of these chemicals must be assessed for emissions from normal operations, *e.g* for determining compliance with OSHA or environmental agency standards. Malfunction of a system, *e.g.*, as recited above, increases such emissions, thereby increasing the risk of operator exposure, and increasing the risk that an instrument may need to be shut down until risk to an operator is reduced and until any regulatory requirements for operation are met.

Emission or leakage of reagents during operation can have consequences beyond risks to personnel and to the environment. As noted above, instruments may need to be removed from operation for cleaning, leading to a temporary decrease in production capacity of a synthesis facility. Further, any emission or leakage may cause damage to parts of the instrument or to other instruments or aspects of the facility, necessitating repair or replacement of any such parts or aspects, increasing the time and cost of bringing an instrument back into operation. Failure to address emissions or leakage concerns may lead to additional expenses for operation of a facility, *e.g.*, costs for increased or improved fire or explosion containment measures, and addition of costs associated with the elimination of any instrument systems or wiring that have not been determined to be safe for use in such hazardous locations (*e.g.*, by reference to controlling codes, such as electrical codes, or codes covering operations in the presence of flammable and combustible liquids).

The synthesizers of the present invention provide a number of novel features that dramatically improve synthesizer performance and safety compared to available synthesizers. These novel features work both independently and in conjunction to provide enhanced performance. For example, the present invention reduces exposure by improving collection and disposal of emissions that occur during the normal operation of various synthesis instruments. In another embodiment, the present invention reduces exposure by improving aspects of the instrument to reduce risk of malfunctions leading to reagent escape from the system, *e.g.*, through leakage, overflow or other spillage.

For example, in some embodiments, the present invention provides a means of collecting emissions from the interior of synthesizers by providing a reagent dispensing station. In one embodiment, the reagent dispensing station is an integral part of the base 2 of the synthesizer, as illustrated in Figures 47A and 47B. In some embodiments, the reagent dispensing station provides an enclosure for collecting emitted gasses. In some embodiments, the enclosure is created by the provision of a panel 73 to enclose a portion of base 2 containing reagent reservoirs 72, as illustrated in Figure 47B. In some embodiments, the panel 73 is movable for easy access to reagent reservoirs. In some embodiments, it is removeably attached. Removable attachment may be accomplished by any suitable means, such as through the use of VELCRO, screws, bolts, pins, magnets, temporary adhesives, and the like. In preferred embodiments, at least a portion of the panel 18 is slidably moveable. In preferred embodiments, at least a portion of panel 18 is transparent. In some embodiments, the enclosure of the reagent dispensing station comprises a viewing window that is not in a panel 73.

In some embodiments, the enclosure comprises ventilation tubing. In preferred embodiments, panel 73 comprises a ventilation port 74, *e.g.*, for attachment to ventilation tubing. Since reagent vapors are typically heavier than air, in preferred embodiments, the ventilation tubing is attached at the bottom for the enclosure. In a particularly preferred embodiment, the ventilation port is positioned toward the rear of the instrument.

In some embodiments, the enclosure further comprises an air inlet. In a preferred embodiment, a clearance 75 between the panel 73 and the base 2 provides an air inlet. In a particularly preferred embodiment, the air inlet is positioned toward the front of the instrument.

The location of the ventilation port 74 and air inlet is not limited to the panel 73. For example, in an alternative embodiment, the reagent dispensing station comprises a stand for holding the reagent bottles and ventilation tubing, wherein the stand holds the reagent reservoirs and the ventilation tubing removes emitted gases.

Ventilation may be continuous or under the control of an operator. For example, in some embodiments, when the panel 73 is in a closed position, ventilation occurs continuously through the ventilation port 74 or at regular intervals. In other embodiments, an operator may manually activate ventilation prior to opening the panel 73. In still other embodiments, ventilation occurs in an automated fashion immediately prior to the opening of panel 73. For example, where the opening of panel 73 is controlled by a computer processor, activation of the "open" routine triggers ventilation prior to the physical opening of panel 73. In still other embodiments, the contents of the reagent containers are monitored by a sensor and the ventilation is triggered when one or more of the reagent containers are depleted. In some embodiments, the panel 73 is also automatically open, indicating the need for additional reagents and/or allowing an automated reagent container delivery system to supply reagents to the system.

In some embodiments, multiwell plates (e.g. 96 well, 384 well, 1536 well, etc) are employed with the synthesizers of the present invention. In certain embodiments, the synthesizers are parts of a full automated process such that oligonucleotides are produced without human interaction. In some embodiments, the oligonucleotides move through the synthesis component, and processing components, on rails.

2. Automated and Fail-Safe Reagent Supply

In some embodiments, the DNA synthesizers in the oligonucleotide synthesis component further comprise an automated reagent supply system. The automated reagent supply system delivers reagents necessary for synthesis to the synthesizers from a central supply area. In some embodiments, the central supply area is provided in an isolated room equipped for accommodating leakage, fires, and explosions without threatening other portions of the synthesis facility, the environment, or humans. Where the central supply area provides reagents for multiple synthesizers, in some embodiments, the system is configured to allow banks of synthesizer or individual synthesizer to be removed from the system (*e.g.*, for maintenance or

repair) without interrupting activity at other synthesizers. Thus, the present invention provides an efficient fail-safe reagent delivery system.

For example, in some embodiments, acetonitrile is supplied via tubing (*e.g.*, stainless steel or TEFLON tubing) through the automated supply system. De-blocking solution may also be supplied directly to DNA synthesizers through tubing. In some preferred embodiments, the reagent supply system tubing is designed to connect directly to the DNA synthesizers without modifying the synthesizers. Additionally, in some embodiments, the central reagent supply is designed to deliver reagents at a constant and controlled pressure. The amount of reagent circulating in the central supply loop is maintained at 8 to 12 times the level needed for synthesis in order to allow standardized pressure at each instrument. The excess reagent also allows new reagent to be added to the system without shutting down. In addition, the excess of reagent allows different types of pressurized reagent containers to be attached to one system. The excess of reagents in one centralized system further allows for one central system for chemical spills and fire suppression.

In some embodiments, the DNA synthesis component includes a centralized argon delivery system. The system includes high-pressure argon tanks adjacent to each bank of synthesizers. These tanks are connected to large, main argon tanks for backup. In some embodiments, the main tanks are run in series. In other embodiments, the main tanks are set up in banks. In some embodiments, the system further includes an automated tank switching system. In some preferred embodiments, the argon delivery system further comprises a tertiary backup system to provide argon in the case of failure of the primary and backup systems.

In some embodiments, one or more branched delivery components are used between the reagent tanks and the individual synthesizers or banks of synthesizers. For example, in some embodiments, acetonitrile is delivered through a branched metal structure (*e.g.*, the structure described in Figure 56). Where more than one branched delivery component is used, in preferred embodiments, each branched delivery component is individually pressurized.

The present invention is not limited by the number of branches in the branched delivery component. In preferred embodiments, each branched delivery component (100) contains ten or more branches (101). Reagent tanks may be connected to the branched delivery components using any number of configurations. For example, in some embodiments, a single reagent tank is

matched with a single branched component. In other embodiments, a plurality of reagent tanks is used to supply reagents to one or more branched components. In some such embodiments, the plurality of tanks may be attached to the branched components through a single feed line, wherein one or a subset of the tanks feeds the branched components until empty (or substantially empty), whereby a second tank or subset of tanks is accessed to maintain a continuous supply of reagent to the one or more branched components. To automate the monitoring and switching of tanks, an ultrasonic level sensor may be applied.

In some embodiments, each branch of the branched delivery component provides reagent to one synthesizer or to a bank of synthesizers through connecting tubing (102). In preferred embodiments, tubing is continuous (*i.e.*, provides a direct connection between the delivery branch and the synthesizer). In some preferred embodiments, the tubing comprises an interior diameter of 0.25 inches or less (*e.g.*, 0.125 inches). In some embodiments, each branch contains one or more valves (preferably one). While the valve may be located at any position along the delivery line, in preferred embodiments, the valve is located in close proximity to the synthesizer. In other embodiments, reagent is provided directly to synthesizers without any joints or valves between the branched delivery component and the synthesizers.

In some embodiments, the solvent is contained in a cabinet designed for the safe storage of flammable chemicals (a "flammables cabinet") and the branched structure is located outside of the cabinet and is fed by the solvent container through tubing passed through the wall of the cabinet. In other embodiments, the reagent and branched system is stored in an explosion proof room or chamber and the solvent is pumped via tubing through the wall of the explosion proof room. In preferred embodiments, all of the tubing from each of the branches is fed through the wall in at a single location (*e.g.*, through a single hole (103) in the wall (104)).

The reagent delivery system of the present invention provides several advantages. For example, such a system allows each synthesizer to be turned off (*e.g.*, for servicing) independent of the other synthesizers. Use of continuous tubing reduces the number of joints and couplings, the areas most vulnerable to failure, between the reagent sources and the synthesizers, thereby reducing the potential for leakage or blockage in the system. Use of continuous tubing through inaccessible or difficult-to-access areas reduces the likelihood that repairs or service will be needed in such areas. In addition, fewer valves results in cost savings.

In some embodiments, the branched tubing structure further provides a sight glass (105). In preferred embodiments, the sight glass is located at the top of the branched delivery structure. The sight glass provides the opportunity for visual and physical sampling of the reagent. For example, in some embodiments, the sight glass includes a sampling valve (106) (*e.g.*, to collect samples for quality control). In some embodiments, the site glass serves as a trap for gas bubbles, to prevent bubbles from entering the connecting tubing (102). In other embodiments, the sight glass contains a vent (*e.g.*, a solenoid valve) for de-gassing of the system (107). In some embodiments, scanning of the sight glass (*e.g.*, spectrophotometrically) and sampling are automated. The automated system provides quality control and feedback (*e.g.*, the presence of contamination).

In other embodiments, the present invention provides a portable reagent delivery system. In some embodiments, the portable reagent delivery system comprises a branched structure connected to solvent tanks that are contained in a flammables cabinet. In preferred embodiments, one reagent delivery system is able to provide sufficient reagent for 40 or more synthesizers. These portable reagent delivery systems of the present invention facilitate the operation of mobile (portable) synthesis facilities. In another embodiment, these portable reagent delivery systems facilitate the operation of flexible synthesis facilities that can be easily re-configured to meet particular needs of individual synthesis projects or contracts. In some embodiments, a synthesis facility comprises multiple portable reagent delivery systems.

3. Waste Collection

In some embodiments, the DNA synthesis component further comprises a centralized waste collection system. The centralized waste collection system comprises cache pots for central waste collection. In some embodiments, the cache pots include level detectors such that when waste level reaches a preset value, a pump is activated to drain the cache into a central collection reservoir. In preferred embodiments, ductwork is provided to gather fumes from cache pots. The fumes are then vented safely through the roof, avoiding exposure of personnel to harmful fumes. In preferred embodiments, the air handling system provides an adequate amount of air exchange per person to ensure that personnel are not exposed to harmful fumes.

The coordinated reagent delivery and waste removal systems increase the safety and health of workers, as well as improving cost savings.

In some embodiments, the solvent waste disposal system comprises a waste transfer system. In some preferred embodiments, the system contains no electronic components. In some preferred embodiments, the system comprises no moving parts. For example, in some embodiments, waste is first collected in a liquid transfer drum (200) designed for the safe storage of flammable waste (See Figure 57 for an exemplary waste disposal system). In some embodiments, waste is manually poured into the drum through a waste channel (201). In preferred embodiments, solvent waste is automatically transported (*e.g.*, through tubing) directly from synthesizers to the drum (200). To drain the liquid transfer drum (200), argon is pumped from a pressurized gas line (202) into the drum through a first opening (203), forcing solvent waste out an output channel (204) at a second opening (205) (*e.g.*, through tubing) into a centralized waste collection area. In preferred embodiments, the argon is pumped at low pressure (*e.g.*, 3-10 pounds per square inch (psi), preferably 5 psi or less). In some embodiments, the drum (200) contains a sight glass (207) to visualize the solvent level. In some embodiments, the level is visualized manually and the disposal system is activated when the drum (200) has reached a selected threshold level (207). In other embodiments, the level is automatically detected and the disposal system is automatically activated when the drum (200) has reached the threshold level (207).

The solvent waste transfer system of the present invention provides several advantages over manual collection and complex systems. The solvent waste system of the present invention is intrinsically safe, as it can be designed with no moving or electrical parts. For example, the system described above is suitable for use in Division I/Class I space under EPA regulations.

Some process steps may put out caustic waste. For example, deprotection of synthesized oligonucleotides generally includes treatment with NH4OH. In some embodiments, caustic waste is neutralized before disposal, *e.g.*, to a sanitary sewer. In preferred embodiments, the neutralization of the waste is checked (*e.g.*, by measurement of pH) to ensure that it is in an appropriate condition for disposal via the intended system (*e.g.*, the sanitary sewer system).

In some embodiments, waste from each deprotection station is neutralized before collection to a centralized waste collection or disposal system. In other embodiments, caustic waste from a plurality of deprotection stations is collected before neutralization.

By way of example, and not intended as a limitation, the following provides a description for one embodiment of a centralized collection and neutralization system for caustic waste. The system may comprise collection of caustic waste from one or more stations in a tank, *e.g.*, a carboy. In some embodiments, the amount of neutralizing reagent required to neutralize a defined amount of caustic waste is calculated, based on the volume and content of the waste. In some embodiments, the calculated amount of neutralizing reagent is added after collection of the waste. In preferred embodiments, the calculated amount of neutralizing reagent is provided in the carboy, such that when the carboy is full or when the combined volume of the neutralizer and waste reaches a predetermined volume, the waste has been neutralized.

In one embodiment, the carboy is provided with a pH probe for measurement of the pH of the collected waste. In some embodiments, the system provides a means of altering the pH of the collected waste. In preferred embodiments, the altering of the pH occurs in response to a measured pH value for the collected waste. For example, if the pH is determined to be outside a certain range, (*e.g.*, if it does not fall between, for example, pH 7 and pH 9), the system provides a reagent selected to adjust the pH to the selected range (*e.g.*, if the pH is found to be high, the system dispenses an acidic solution for neutralization; if the pH is low, the system dispenses a basic solution for neutralization). When the pH comes into the selected range, the system shuts off the dispenser. For the step of dispensing a neutralizing reagent, any system suitable for the controlled delivery of a reagent is contemplated. For example, discharge may be accomplished via a mechanical dispenser, or discharge can be accomplished via non-mechanical means, *e.g.*, via control of air pressure.

In some embodiments, neutralization treatment is provided to the collected waste in bulk, *e.g.*, when the carboy is full or when it reaches a predetermined threshold level. In other embodiments, neutralization is periodic. In some embodiments, periodic neutralization is set to occur at particular times, *e.g.*, at particular times of day, or whenever a particular interval of time has passed since the last treatment. In other embodiments, periodic treatment is set to respond to a condition of the waste container, such as whenever a new addition of waste material occurs, or

whenever the pH is not within the selected range. In yet other embodiments, periodic treatment occurs based on a combination of these or other factors.

In a preferred embodiment, the carboy is provided with a means for mixing, such as a stirrer or agitator. In some embodiments, the system comprises a device for keeping a precipitate suspended. In some embodiments, the system provides a filter for removing precipitates, particulates or other non-liquid matter in the collected waste. In other preferred embodiments, the system provides a means of venting gasses. In particularly preferred embodiments, the gasses are collected for disposal through a centralized ventilation system.

4. Centralized Control System

In some embodiments, all of the DNA synthesizers in the synthesis component are attached to a centralized control system. The centralized control system controls all areas of operation, including, but not limited to, power, pressure, reagent delivery, waste, and synthesis. In preferred embodiments, the centralized control system is operably linked to data (enterprise) management system (See, below). In other preferred embodiments, the centralized control system (for oligonucleotide synthesis) is operably linked to the centralized control network (for oligonucleotide processing. The combination of the centralized control system and centralized control network is referred to as the shop floor control system. In some preferred embodiments, the centralized control system includes a clean electrical grid with uninterrupted power supply. Such a system minimizes power level fluctuations. In additional preferred embodiments, the centralized control system includes alarms for air flow, status of reagents, and status of waste containers. The alarm system can be monitored from the central control panel. The centralized control system allows additions, deletions, or shutdowns of one synthesizer or one block of synthesizers without disrupting operations of other instruments. The centralized power control allows user to turn instruments off instrument by instrument, bank by bank, or the entire module. In some embodiments, the centralized control system comprises enterprise software (e.g. Oracle, PeopleSoft, etc.).

B. Oligonucleotide Processing Components

In some embodiments, the automated DNA production process further comprises one or more oligonucleotide production components, including, but not limited to, an oligonucleotide cleavage and deprotection component, an oligonucleotide purification component, a dry-down component, a desalting component, a dilution and fill component, and a quality control component. In preferred embodiments, the synthesis component is integrated with the oligonucleotide processing components, and other components such as the order entry component discussed above (see also Figure 58b). Preferably, the components are operably linked for data sharing, product tracking and control. It is also preferred that the various components are operably linked such that oligonucleotides are processed with limited human interaction. A general overview of how the components are operably connected, in some embodiments, is provided in Figure 58a. Particular embodiments for process and data flow within and between the various processing components are shown in Figures 58b-58k.

Preferably the oligonucleotide components are automated, at least in part, in order to improve efficiencies and reduce human errors. In preferred embodiments, 96 well (or 384 well) plates are used through out the entire system (e.g. from initial synthesis to dilute and fill), such that individual columns do not have to be transferred between different sized plates. In other embodiments, samples are maintained in a closed-circuit tubing for synthesis and one or more additional components (e.g., cleavage and deprotection, purification, etc.) such that a solution carrying the sample passes through a plurality of reaction zones where the tubing is heated, agitated, accessed by other tubing to deliver necessary reagents, etc. without ever being removed from the tubing or exposed to the ambient environment. Such systems facilitate high-throughput production if detection assays.

1. Oligonucleotide Cleavage and Deprotection

After synthesis is complete, the oligonucleotides synthesis columns are moved to the cleavage and deprotection station. In some embodiments, the transfer of oligonucleotides to this station is automated and controlled by robotic automation. In some embodiments, the entire cleavage and deprotection process is performed by robotic automation. In some embodiments, NH_4OH for deprotection is supplied through the automated reagent supply system.

Accordingly, in some embodiments, oligonucleotide deprotection is performed in multi-sample containers (e.g., 96 well covered dishes) in an oven. This method is designed for the high-throughput system of the present invention and is capable of the simultaneous processing of large numbers of samples. This method provides several advantages over the standard method of deprotection in vials. For example, sample handling is reduced (*e.g.*, labeling of vials dispensing of concentrated NH_4OH to individual vials, as well as the associated capping and uncapping of the vials, is eliminated). This reduces the risks of contamination or mislabeling and decreases processing time. Where such methods are used to replace human pipetting of samples and capping of vials, the methods save many labor hours per day. The method also reduces consumable requirements by eliminating the need for vials and pipette tips, reduces equipment needs by eliminating the need for pipettes, and improves worker safety conditions by reducing worker exposure to ammonium hydroxide. The potential for repetitive motion disorders is also reduced. Deprotection in a multi-well plate further has the advantage that the plate can be directly placed on an automated desalting apparatus (*e.g.*, TECAN Robot).

During the development of the present invention, the plate was optimized to be functional and compatible with the deprotection methods. In some embodiments, the plate is designed to be able to hold as much as two milliliters of oligonucleotide and ammonium hydroxide. If deep well plates are used, automated downstream processing steps may need to be altered to ensure that the full volume of sample is extracted from the wells. In some embodiments, the multi-well plates used in the methods of the present invention comprise a tight sealing lid/cover to protect from evaporation, provide for even heating, and are able to withstand temperatures and pressures necessary for deprotection. Attempts with initial plates were not successful, having problems with lids that were not suitably sealed and plates that did not withstand deprotection temperatures.

In some embodiments (*e.g.*, processing of target and INVADER oligonucleotides), oligonucleotides are cleaved from the synthesis support in the multi-well plates. In other embodiments (*e.g.*, processing of probe oligonucleotides), oligonucleotides are first cleaved from the synthesis column and then transferred to the plate for deprotection.

In preferred embodiments, the present invention provides devices and systems for automated and semi-automated cleavage and/or protections. Preferably, the cleave and deprotect

device is configured to hold 96 synthesis columns (e.g. in an 8 by 12 plate). It is also preferred that reagents, such as ammonium hydroxide, may be contacted with the synthesis columns (or other columns containing oligonucleotides) with minimal or no exposure of the reagents to the ambient environment. Also, the cleave and deprotect device is preferably configured to allow the automatic dispersement of reagents into the synthesis columns at periodic intervals in order to facilitate cleavage. For example, the present invention provides a system comprising a series of fluid dispensers (e.g. a series of fluid dispensers), a software application (e.g. Unicorn software) that instructs the fluid dispenser (e.g. to engage the synthesis columns once the rack holding the columns is inserted into the automated device), and a cleave and deprotect device for holding the synthesis columns. In other preferred embodiments, the cleave and deprotect device allows reagents such as ammonium hydroxide to pass through the synthesis column and into a receive plate below (e.g. a 96 well receive plate that collects the reagents and oligonucleotdies as they are cleaved from the synthesis columns). The receiving plate may be in a 96 well, 384 well, or any other type of format. In other preferred embodiments, the fluid is dispensed in lines that end with fluid column connections (e.g. Figure 60A, number 106), or the fluid column connections are part of the cleave and deprotect device.

Figure 60 shows exemplary components of an automated cleave and deprotect system. Figures 60A and 60B show a side view of a cleave and deprotect device. Figure 60A shows the fluid column connections in the down position (e.g. engaged with the synthesis columns), and Figure 60B shows the fluid column connections in the up position. A brief description of various part of the cleavage and/or deprotect device as shown in Figures 60A-H is provided. The catch plate 100 is preferably a deep well plate. This catch plate collects the oligonucleotides as they come off the column due to exposure to ammonium hydroxide. The catch plate may, for example, be a 96 well plate. This plate can them be moved to a further processing step (e.g. a deprotection step, where the plate is covered and then heat is applied). Columns 102 (e.g. synthesis columns) are held in column holder 104 (See Figure 60A). A top view of one particular column holder is provided in Figure 60E. Fluid column connection 106 allows liquid to be dispensed to the columns with minimal or no exposure of reagents to the ambient environment. Fluid column connections may be made from any suitable material, and have various parts that facilitate connection with the columns (see Figure 60F). Connection 106 has a

plurality of rings 108 (2 shown in Figure 60A). Either one or both rings engage the interior surface 10 of column. The rings 108 are radiused so that they form a releasable seal whey they engage surface 110. It is appreciated that when rings 108 are radiused a releasable seal is formed even if columns 108 are at an angle other than a 9 degree angle to column holder 104. Even if there is a small amount of misalignment between the column 102 and connection 106 there is a substantially airtight and water tight seal formed.

Columns 102 when releasably sealed to connections 106 move horizontally and/or vertically as a block in some embodiments. When the columns 102 rise up with connections they contact stripper plate 112 which has an aperature 114 which permits connection 106 to pass therethrough, but acts as a limit stop when lip 118 contacts stripper block plate surface 120 (see Stripper plate in Figure 60A and Figure 60C). Aperature 114 is large enough to let the connection 106 to ride through it but is smaller than the diameter of lip 118. Actuation of connection holder 122 for movement along the guide shafts 124 (see Figures 60A and 60H) which are secured to base 126. The base of the machine is shown in Figures 60A and 60G. Finally the dispense tip holder is shown in Figures 60A and 60D.

In some embodiments, software, such as Unicorn Software, controls the amount and timing of reagents dispensed into the synthesis columns. For example, a 45 minute program may be run that periodically dispenses ammonium hydroxide into the synthesis columns at timed intervals in order to cleave the oligonucleotides off of the synthesis columns. In certain embodiments, the automated cleavage and deprotection system is configured to work with a polyplex machine (e.g. software allows an interface between the cleavage and deprotection).

In certain embodiments, fast deprotection chemistry is utilized to increase the rate at which oligonucleotide are manufactured. For example, oligonucleotdies may be synthesized with Proligo Tac Amidites that have a tert.-butylphenoxy-acetyl "tac" base protecting group. This protecting group decreases cleavage and deprotection time of the final oligo from about eight hours to about 15 minutes at 55 °C, or two hours at room temperature when compared with standard base protecting groups. Rapid deprotection results in less exposure to ammonia and reduced risk of hydrolysis. Also, this type of fast deprotection chemistry may be used with the autocleave device of the present invention. For example, the autocleave device may be heated up to the deprotecting temperature (e.g. 60 degrees Celsius), and both cleavage and deprotection

can occur in the same column in the autocleave device. This allows, for example, the cleaved and deprotected to go straight into a purification column (e.g. C_{18} column).

2. Oligonucleotide Purification

In some embodiments, following deprotection and cleavage from the solid support, oligonucleotides are further purified. In certain embodiments, the purification step is not necessary (e.g. the synthesis and cleave and deprotect steps yield a sufficiently pure oligonucleotide preparation, or the detection assay being produced does not require an oligonucleotide purification step). Any suitable purification method may be employed when purification is desired, including, but not limited to, high pressure liquid chromatography (HPLC) (e.g., using reverse phase C18 and ion exchange), reverse phase cartridge purification, probe capture, and gel electrophoresis. However, in preferred embodiments, purification is carried out using ion exchange HPLC chromatography.

In some embodiments, multiple HPLC instruments are utilized, and integrated into banks (*e.g.*, banks of 8 HPLC instruments). Each bank is referred to as an HPLC module. Each HPLC module consists of an automated injector (*e.g.*, including, but not limited to, Leap Technologies 8-port injector) connected to each bank of automated HPLC instruments (*e.g.*, including, but not limited to, Beckman-Coulter HPLC instruments). The automatic Leap injector can handle four 96-well plates of cleaved and deprotected oligonucleotides at a time. The Leap injector automatically loads a sample onto each of the HPLCs in a given bank. The use of one injector with each bank of HPLC provides the advantage of reducing labor and allowing integrated processing of information. In preferred embodiments, reagents are supplied directly to the HPLC instruments via a solvent delivery component (See, e.g. Figure 56).

In some embodiments, oligonucleotides are purified on an ion exchange column using a salt gradient. Any suitable ion exchange functionality or support may be utilized, including but not limited to, Source 15 Q ion exchange resin (Pharmacia). Any suitable salt may be utilized for elution of oligonucleotides from the ion exchange column, including but not limited to, sodium chloride, acetonitrile, and sodium perchlorate. However, in preferred embodiments, a gradient of sodium perchlorate in acetonitrile and sodium acetate is utilized.

In some embodiments, the gradient is run for a sufficient time course to capture a broad range of sizes of oligonucleotides. For example, in some embodiments, the gradient is a 54 minute gradient carried out using the method described in Tables 3 and 4. Table 3 describes the HPLC protocol for the gradient. The time column represents the time of the operation. The module column represents the equipment that controls the operation. The function column represents the function that the HPLC is performing. The value column represents the value of the HPLC function at the time specified in the time column. Table 4 describes the gradient used in HPLC purification. The column temperature is approximately 65°C. Buffer A is 20 mM Sodium Perchlorate, 20 mM Sodium Acetate, 10 percent Acetonitrile, pH 7.35. Buffer B is 600 mM Sodium Perchlorate, 20 mM Sodium Acetate, 10 percent Acetonitrile, pH 7-8.

In some embodiments, the gradient is shortened. In preferred embodiments, the gradient is shortened so that a particular gradient range suitable for the elution of a particular oligonucleotide being purified is accomplished in a reduced amount of time. In other preferred embodiments, the gradient is shortened so that a particular gradient range suitable for the elution of any oligonucleotide having a size within a selected size range is accomplished in a reduced amount of time. This latter embodiment provides the advantages that the worker performing HPLC need not have foreknowledge of the size of an oligonucleotide within the selected size range, and the protocol need not be altered for purification of any oligonucleotide having a size within the range.

In a particularly preferred embodiment, the gradient is a 34 minute gradient described in the Tables 4 and 5. The parameters and buffer compositions are as described for Tables 3 and 4 above. Reducing the gradient to 34 minutes increases the capacity of synthesis per HPLC instrument and reduces buffer usage by 50% compared to the 54 minute protocol described above. The 34 minute HPLC method of the present invention has the further advantage of being optimized to be able to separate oligonucleotides of a length range of 23-39 nucleotides without any changes in the protocol for the different lengths within the range. Previous methods required changes for every 2-3 nucleotide change in length. In yet other embodiments, the gradient time is reduced even further (*e.g.*, to less than 30 minutes, preferably to less than 20 minutes, and even more preferably, to less than 15 minutes). Any suitable method may be utilized that meets the

requirements of the present invention (*e.g.*, able to purify a wide range of oligonucleotide lengths using the same protocol).

In some embodiments, separate sets of HPLC conditions, each selected to purify oligonucleotides within a different size range, may be provided (*e.g.*, may be run on separate HPLCs or banks of HPLCs). Thus, in some embodiments of the present invention, a first bank of HPLCs are configured to purify oligonucleotides using a first set of purification conditions (e.g., for 23-39 mers), while second and third banks are used for the shorter and longer oligonucleotides. Use of this system allows for automated purification without the need to change any parameters from purification to purification and decreases the time required for oligonucleotide production.

In some embodiments, the HPLC station is equipped with a central reagent supply system. In some embodiments, the central reagent system includes an automated buffer preparation system. The automated buffer preparation system includes large vat carboys that receive pre-measured reagents and water for centralized buffer preparation. The buffers (*e.g.*, a high salt buffer and a low salt buffer) are piped through a circulation loop directly from the central preparation area to the HPLCs. In some embodiments, the conductivity of the solution in the circulation loop is monitored to verify correct content and adequate mixing. In addition, in some embodiments, circulation lines are fitted with venturis for static mixing of the solutions as they are circulated through the piping loop. In still further embodiments, the circulation lines are fitted with 0.05 μm filters for sterilization.

In some preferred embodiments, the HPLC purification step is carried out in a clean room environment. The clean room includes a HEPA filtration system. All personnel in the clean room are outfitted with protective gloves, hair coverings, and foot coverings.

In preferred embodiments, the automated buffer prep system is located in a non- clean room environment and the prepared buffer is piped through the wall into the clean room.

Each purified oligonucleotide is collected into a tube (*e.g.*, a 50-ml conical tube) in a carrying case in the fraction collector. Collection is based on a set method, which is triggered by an absorbance rate change, level, or threshold within a predetermined time window. In some embodiments, the method uses a flow rate of 5 ml/min (the maximum rate of the pumps is 10 ml/min.) and each column is automatically washed before the injector loads the next sample.

(Det = detector; %B = percent of buffer B; flow rate values in ml/min)

Table 3

54 Minute HPLC Method

Time (min)	Module	Function	Value	Duration (min)
0	Pump	%B	22.00	4.0
0	Det 166-3	Autozero ON		
0	Det 166-3	Relay ON	3.0	0.10
4	Pump	%B	37.00	43.00
47	Pump	%B	100.00	0.50
47.5	Pump	Flow Rate	7.5	0.00
50.0	Pump	%B	5.0	0.50
53.45	Det 166-3	Stop Data		

Table 4

54 Minute HPLC Method

Time	Gradient	Flow Rate
0	5% B/ 95% A	5 ml/min
0-4 min	5-22% B	5 ml/min
4-47 min	22-37% B	5 ml/min
47-47.5 min	37-100% B	7.5 ml/min
47.5-50 min	100% B	7.5 ml/min
50-50.5 min	100-5% B	7.5 ml/min
50.5-53.5 min	5% B	7.5 ml/min

Table 5

34 Minute HPLC Method

Time (min)	Module	Function	Value	Duration
0	Pump	%B	26.00	2.0

0	Det 166-3	Autozero ON		
0	Det 166-3	Relay ON	3.0	0.10
2	Pump	%B	36.00	27.00
29	Pump	%B	100.00	0.50
29.5	Pump	Flow Rate	7.5	0.00
32	Pump	%B	5.0	0.50
33.45	Det 166-3	Stop Data		

Table 6 34 Minute HPLC Method		
Time	Gradient	Flow Rate
0	5% B/ 95% A	5 ml/min
0-2 min	5-26% B	5 ml/min
2-29 min	26-36% B	5 ml/min
29-29.5 min	36-100% B	6.5 ml/min
29.5-32 min	100% B	7.5 ml/min
32-32.5 min	100-5% B	7.5 ml/min
32.5-33.5 min	5% B	7.5 ml/min

3. Dry-Down Component

When the fraction collector is full of eluted oligonucleotides, they are transferred (*e.g.*, by automated robotics or by hand) to a drying station. For example, in some embodiments, the samples are transferred to customized racks for Genevac centrifugal evaporator to be dried down. In preferred embodiments, the Genevac evaporator is equipped with racks designed to be used in both the Genevac and the subsequent desalting step. The Genevac evaporator decreases drying time, relative to other commercially available evaporators, by 60%.

4. Desalting Component

In some embodiments, following HPLC, oligonucleotides are desalted. In other embodiments, oligonucleotides are not HPLC purified, but instead proceed directly from deprotection to desalting. In some embodiments, the desalting stations have TECAN robot systems for automated desalting. The system employs a rack that has been designed to fit the TECAN robot and the Genevac centrifugal evaporator without transfer to a different rack or holder. The racks are designed to hold the different sizes of desalting columns, such as the NAP-5 and NAP-10 columns. The TECAN robot loads each oligonucleotide onto an individual NAP-5 or NAP-10 column, supplies the buffer, and collects the eluate. If desired, desalted oligonucleotides may be frozen or dried down at this point.

In some embodiments, following desalting, INVADER and target oligonucleotides are analyzed by mass spectroscopy. For example, in some embodiments, a small sample from the desalted oligonucleotide sample is removed (e.g., by a TECAN robot) and spotted on an analysis plate, which is then placed into a mass spectrometer. The results are analyzed and processed by a software routine. Following the analysis, failed oligonucleotides are automatically reordered, while oligonucleotides that pass the analysis are transported to the next processing step. This preliminary quality control analysis removes failed oligonucleotides earlier in the processing, thus resulting in cost savings and improving cycle times.

5. Oligonucleotide Dilution and Fill Component

In some embodiments, the oligonucleotide production process further includes a dilute and fill module. In some embodiments, each module consists of three automated oligonucleotide dilution and normalization stations. Each station consists of a network-linked computer and an automated robotic system (*e.g.*, including but not limited to Biomek 2000). In one embodiment, the pipetting station is physically integrated with a spectrophotometer to allow machine handling of every step in the process. All manipulations are carried out in a HEPA-filtered environment. Dissolved oligonucleotides are loaded onto the Biomek 2000 deck the sequence files are transferred into the Biomek 2000. The Biomek 2000 automatically transfers a sample of each oligonucleotide to an optical plate, which the spectrophotometer reads to measure the A260 absorbance. Once the A260 has been determined, an Excel program integrated with the Biomek

software uses absorbance and the sequence information to prepare a dilution table for each oligonucleotide. The Biomek employs that dilution table to dilute each oligonucleotide appropriately. The instrument then dispenses oligonucleotides into an appropriate vessel (*e.g.*, 1.5 ml microtubes).

In some preferred embodiments, the automated dilution and fill system is able to dilute different components of a kit (*e.g.*, INVADER and probe oligonucleotides) to different concentrations. In other preferred embodiments, the automated dilution and fill module is able to dilute different components to different concentrations specified by the end user.

6. Quality Control Component

In some embodiments, oligonucleotides undergo a quality control assay before distribution to the user. The specific quality control assay chosen depends on the final use of the oligonucleotides. For example, if the oligonucleotides are to be used in an INVADER SNP detection assay, they are tested in the assay before distribution.

In some embodiments, each SNP set is tested in a quality control assay utilizing the Beckman Coulter SAGIAN CORE System. In some embodiments, the results are read on a real-time instrument (*e.g.*, a ABI 7700 fluorescence reader). The QC assay uses two no target blanks as negative controls and five untyped genomic samples as targets. For consistency, every SNP set is tested with the same genomic samples. In preferred embodiment, the ADS system is responsible for tracking tubes through the QC module. Thus, in some embodiments, if a tube is missing, the ADS program discards, reorders, or searches for the missing tube.

In some preferred embodiments, the user chooses which QC method to run. The operator then chooses how many sets are needed. Then, in some embodiments, the application auto-selects the correct number of SNPs based on priority and prints output (picklist). If a picklist needs to be regenerated, the operator inputs which picklist they are replacing as well as which sets are not valid. The system auto-selects the valid SNPs plus replacement SNPs and print output. Additionally, in some embodiments, picklists are manually generated by SNP number.

The auto-selected SNPs are then removed from being listed as available for auto-selection. In some embodiments, the software prints the following items: SNP/Oligo list

(picklist), SNP/Oligo layout (rack setup). The operator then takes the picklist into inventory and removes the completed oligonucleotide sets. In some embodiments, a completed set is unavailable. In this case, the operator regenerates a picklist. Then, in preferred embodiments, the missing SNP set or tube is flagged in the system. Once a picklist is full, the oligonucleotides are moved to the next step.

In some embodiments, the operator then takes the rack setup generated by the picklist and loads the rack. Alternatively, a robotic handling system loads the rack. In preferred embodiments, tubes are scanned as they are placed onto the rack. The scan checks to make sure it is the correct tube and displays the location in the rack where the tube is to be placed.

Completed racks are then placed in a holding area to await the robot prep and robot run. Then, in some embodiments, the operator views what racks are in the queue and determines what genomics and reagent stock will be loaded onto the robot. The robot is then programmed to perform a specific method. Additionally, in some embodiments, the robot or operator records genomics and reagents lot numbers.

In preferred embodiments, a carousel location map is printed that outlines where racks are to be placed. The operator then loads the robot carousel according to the method layout. The rack is scanned (*e.g.*, by the operator or by the ADS program).. If the rack is not valid for the current robot method, the operator will be informed. The carousel location for the rack is then displayed. The output plates are then scanned (*e.g.*, by the operator or by the ADS program). If the plate is not valid for the current method the operator is informed. The carousel location for the plate is then displayed.

Then, in some embodiments, the robot is run. The robot then places the plates onto heatblocks for a period of time specified in the method. In some embodiments, the robot then scans the plates on the Cytofluor. Output from the cytofluor is read into the database and attached to the output plate record.

In other embodiments, the output is read on the ABI 7700 real time instrument. In some embodiments, the operator loads the plate on to the 7700. Alternatively, in other embodiments, the robot loads the plate onto the ABI 7700. A scan is then started using the 7700 software. When the scan is completed the output file is saved onto a computer hard drive. The operator then starts the application and scans in the plate bar code. The software instructs the user to

browse to the saved output file. The software then reads the file into the database and deletes the file (or tells the operator to delete the file).

The plate reader results (*e.g.,* from a Cytofluor or a ABI 7700) are then analyzed (*e.g.*, by a software program or by the operator). The present invention provides assessment methods to determine if a particular detection assay will pass the quality control component. The assessment process reviews the performance of the manufactured components (oligos: probe, invader, synthetic targets and CLEAVASE enzyme) of the detection assay (e.g., INVADER Assay, TAQMAN assay, etc.) under conditions similar, if not identical, to those that will be used by the customer. This automated process produces an assessment result ("PASS" or "FAIL") and instructions as to the disposition (e.g. keep, reorder, resynthesize, bin) of the component oligonucleotides (ODNs) (e.g., probes, invader, targets) comprising the Assay. The latter role, the automated production of ODN disposition instructions, is an integral part of the overall modular and automated ODN production process due to the numerous platforms and configurations under which the INVADER Assay can be utilized.

This is achieved, for example, by testing an assay against several target types or classes, such as: No Target, Synthetic Target and Genomic Target. Utilizing these classes allows for the assessment process to be broken down into modules allowing for the numerous data and derived performance metrics to be funneled into an overall singular Pass/Fail code with the corresponding instructions for the disposition of the assay components.

This process may be employed, for example, for the assessment of the ODN components comprising the INVADER Assay. However, the assessment process may also be applied to the assessment of other assays (e.g. TAQMAN) and the ODN components that comprise other types of detection assays.

The assessment process of the present invention may be carried out in a series of steps.

Step 1 – Assay format

The assay format is based on the number of targets within each class is to be tested as well as the number of repetitions to which each target will be subjected.

Step 2 – Allele Call process

The general process for step 2 is outlined in Figure 97A. In the case of a biplex assay, an allele call/identification may be made by analyzing the raw data to derive three performance

metrics, the FOZ (fold over zero) (calculated per signal dye/allele), and a FOZ Ratio. These metrics are compared to minimal threshold levels for making a genotyping call (Heterozygous, $Homozygous_{WT}$, $Homozygous_{Mut}$, or Equivocal/Ambiguous). If the two FOZ values can make a genotyping call that agrees with one made by the FOZ Ratio then the allele call is validated. Both validated calls and invalidated calls are then coded.

Performance metrics

Performance metrics are those values that are mathematically derived from the raw data. The raw data is that generated by the device/instrument used to measure the assay performance (real-time or endpoint mode).

<u>FOZ or S/NT</u>

$FOZ_{Dye1} = (RawSignal_{Dye1}/NTC_{Dye1})$

$FOZ_{Dye2} = (RawSignal_{Dye2}/NTC_{Dye2})$

In the case of replicated runs, $RawSignal_{DyeX}$ and NTC_{DyeX} are the averaged values.

<u>FOZ Ratio</u>

$\text{FOZ Ratio} = (1- FOZ_{Dye1}) / (1- FOZ_{Dye2})$

<u>CV</u>

$\text{Coefficient of Variance} = StDev_{signal}/Avg_{signal}$

Performance codes

Performance codes are those values that are generated based on the comparison of the aforementioned performance metrics to threshold metric values. This codification step not only sets the minimal metric value that can be used for making allele calls, but it also codifies why a specific well's performance metric failed.

Step 3 – Class Analysis

The general process for step 3 is outlined in Figure 97B. Allele Calls, both valid and invalid are grouped according to the target class, either genomic or synthetic. Each well's calls are then sorted into two cases, valid and invalid calls.

Case 1: Valid Calls

Valid calls are simply tallied as either Homozygous (WT or Mut) and Heterozygous. Note that depending on the assay format/formulation, a Heterozygous call for synthetic targets may be deemed an invalid call.

Case 2: Invalid Calls

Invalid calls are those in which the genotype called using FOZs do not agree with those called using the FOZ Ratio method. Invalid calls may then be analyzed, depending on what target class, using a Failure Metrix that identifies the failing component ODN.

A Class Analysis Code is then generated by tallying the number of valid calls, sorted by genotype, and invalid calls, sorted by component ODN failure.

Step 4 – Class Pass/Fail Flag

The general procedure for step 4 is outlined in Figure 97C. The Class Analysis Codes are used and screened against a set of pass/fail/retest criteria which include:

Minimum number of Valid Calls – unambiguous or equivocal calls count against this number.

Allele representation – P/F/R (Pass/Fail/Retest) for the target class is based on a minimum number of Valid Homozygous calls for each allele that must be present in the tested target population.

Reproducibility – as reflected in the threshold CV value.

Step 5 – SNP P/F/R

The general procedure for step 5 is presented in Figure 97D. The status of the current SNP component ODNs is determined by the comparison/classification of the determined Class P/F/R Flag and the Class Analysis Codes. Weighting of one class over the other may be varied and is dependent upon the QC specification per customer and/or format. Recommendations as to the overall failure status of a particular component ODN may change depending on the result of another target Class Analysis Code and Class P/F/R Flag. A final SNP PFCode is issued which includes the total number of valid calls and the number of times a component ODN was deemed a failure.

Step 6 – Component ODN disposition

The general procedure for step 6 (and step 5) is presented in Figure 97D. Depending on the result of the *SNP PFCode* the current SNP component ODN package is classified into the categories:

PASS

The component ODNs are all marked for shipment and the recommendation is forwarded to the appropriate production module.

FAIL

Instructions as to the disposition of each of the component ODNs are determined from the *SNP PFCode*. An action code is issued and is sent to the to appropriate production modules for processing (resynthesis/reorder).

RETEST

The component ODNs are saved and returned to the queue for retesting (not resythesized or reordered)

In some embodiments, the operator reviews the results of the software analysis of each SNP and takes one of several actions. In some embodiments, the operator approves all automated actions. In other embodiments, the operator reviews and approves individual actions. In some embodiments, the operator marks actions as needing additional review. Alternatively, in other embodiments, the operator passes on reviewing anything. Additionally, in some embodiments, the operator overrides all automated actions.

Depending on the results of the QC analysis, one of several actions is next taken. If the software marks ready for Full Fill, the operator forwards discards diluted Probe/INVADER oligonucleotide mixes and forwards the samples to the packaging module.

If an oligonucleotide set fails quality control, the data is interpreted to determine the cause of the failure. The course of action is determined by such data interpretation. If the software marks an oligonucleotide Reassess Failed Oligonucleotide, no action by user is required, the reassess is handled by automation. In the software marks an oligonucleotide Redilute Failed Oligonucleotide, the operator discards diluted tubes. No other action is required. If the software marks an oligonucleotide Order Target Oligonucleotide, no action by user is required. In this case, a synthetic target oligonucleotide is ordered for further testing. If the

software marks an oligonucleotide Fail Oligo(s) Discard Oligo(s), the operator discards the diluted tubes and un-diluted tubes. No other action is required. If the software marks an oligonucleotide Fail SNP, the operator discards the diluted and un-diluted tubes. No other action is required. If the software marks an oligonucleotide Full SNP Redesign, the operator discards the diluted and un-diluted tubes. No other action is required. If the software marks an oligonucleotide Partial SNP Redesign the operator discards diluted tubes and discards some un-diluted tubes. No other action is required.

In some embodiments, the software marks an oligonucleotide Manual Intervention. This step occurs if the operator or software has determined the SNP requires manual attention. This step puts the SNP "on hold" in the tracking system while the operator investigates the source of the failure.

When a set of oligonucleotides (*e.g.*, a INVADER assay set) is completed, the set is transferred to the packaging station.

In some embodiments of the present invention, the produced detection assays are tested against a plurality of samples representing two or more different alleles (samples containing sequences from individuals with different ethnic backgrounds, disease states, etc.) to demonstrate the viability of the assay with different individuals. In preferred embodiments, the produced assays are tested against a sufficient number of alleles (e.g., 100 or more) to identify which members of the population can be tested by the assay and to identify the allele frequency in the population of the genotype for which the assay is designed. In some embodiments, where certain individuals or classes of individuals are not detected by the detection assay, the target sequence of the individuals is characterized to determine whether the intended SNP is not present and/or whether additional mutations are present the prevent the proper detection of the sample. Any such information may be collected and stored in databases. In some embodiments, target selection, in silico analysis, and oligonucleotide design are repeated to generate assays capable of detecting the corresponding sequence of these individuals, as desired. In some embodiments, allele frequency information is stored in a database and made available to users of the detection assays upon request (e.g., made available over a communication network).

C. Packaging Component

In some embodiments, one or more components generated using the system of the present invention are packaged using any suitable means. In some embodiments, the packaging system is automated. In some embodiments, the packaging component is controlled by the centralized control network of the present invention.

D. Centralized Control Network

In some embodiments, the automated DNA production process further comprises a centralized control system. In some embodiments, the centralized control system comprises a computer system. In preferred embodiments, the centralized control system is operably linked to data (enterprise) management system (See, below). Figure 58a-58k shows how the centralized control network if configured in some embodiments of the present invention.

In preferred embodiments, the centralized control network (for oligonucleotide processing) is operably linked to the centralized control system (for oligonucleotide synthesis). The combination of the centralized control system and centralized control network is referred to as the shop floor control system.

In some embodiments, the computer system comprises computer memory or a computer memory device and a computer processor. In some embodiments, the computer memory (or computer memory device) and computer processor are part of the same computer. In other embodiments, the computer memory device or computer memory are located on one computer and the computer processor is located on a different computer. In some embodiments, the computer memory is connected to the computer processor through the Internet or World Wide Web. In some embodiments, the computer memory is on a computer readable medium (*e.g.*, floppy disk, hard disk, compact disk, DVD, etc). In other embodiments, the computer memory (or computer memory device) and computer processor are connected via a local network or intranet. In certain embodiments, the computer system comprises a computer memory device, a computer processor, an interactive device (*e.g.*, keyboard, mouse, voice recognition system), and a display system (*e.g.*, monitor, speaker system, etc.).

In preferred embodiments, the systems and methods of the present invention comprise a centralized control system, wherein the centralized control system comprises a computer tracking

system. As discussed above, the items to be manufactured (*e.g.* oligonucleotide probes, targets, etc) are subjected to a number of processing steps (*e.g.* synthesis, purification, quality control, etc). Also as discussed above, various components of a single order (*e.g.* one type of SNP detection kit) may be manufactured in separate tubes, and may be subjected to a different number of processing steps. Consequently, the present invention provides systems and methods for tracking the location and status of the items to be manufactured such that multiple components of a single order can be separately manufactured and brought back together at the appropriate time. The tracking system and methods of the present invention also allow for increased quality control and production efficiency.

In some embodiments, the computer tracking system comprises a central processing unit (CPU) and a central database. The central database is the central repository of information about manufacturing orders that are received (*e.g.* SNP sequence to be detected, final dilution requirements, etc), as well as manufacturing orders that have been processed (*e.g.* processed by software applications that determine optimal nucleic acid sequences, and applications that assign unique identifiers to orders). Manufacturing orders that have been processed may generate, for example, the number and types of oligonucleotides that need to be manufactured (*e.g.* probe, INVADER oligonucleotide, synthetic target), and the unique identifier associated with the entire order as well as unique identifiers for each component of an order (*e.g.* probe, INVADER oligonucleotide, etc). In certain embodiments, the components of an order proceed through the manufacturing process in containers that have been labeled with unique identifiers (*e.g.* bar coded test tubes, color coded test tubes, etc.).

In certain embodiments, the computer tracking system further comprises one or more scanning units capable of reading the unique identifier associated with each labeled container. In some embodiments, the scanning units are portable (*e.g.* hand held scanner employed by an operator to scan a labeled container). In other embodiments, the scanning units are stationary (*e.g.* built into each module). In some embodiments, at least one scanning unit is portable and at least one scanning unit is stationary (*e.g.* hand held human implemented device).

Stationary scanning units may, for example, collect information from the unique identifier on a labeled container (*i.e.* the labeled container is 'red') as it passes through part of one of the production modules. For example, a rack of 100 labeled containers may pass from the

purification module to the dilute and fill module on a conveyor belt or other transport means, and the 100 labeled containers may be read by the stationary scanning unit. Likewise, a portable scanning unit may be employed to collect the information from the labeled containers as they pass from one production module to the next, or at different points within a production module. The scanning units may also be employed, for example, to determine the identity of a labeled container that has been tested (*e.g.* concentration of sample inside container is tested and the identity of the container is determined).

The scanning units are capable of transmitting the information they collect from the labeled containers to a central database. The scanning units may be linked to a central database via wires, or the information may be transmitted to the central database. The central database collects and processes this information such that the location and status of individual orders and components of orders can be tracked (*e.g.* information about when the order is likely to complete the manufacturing process may be obtained from the system). The central database also collects information from any type of sample analysis performed within each module (*e.g.* concentration measurements made during dilute and fill module). This sample analysis is correlated with the unique identifiers on each labeled container such that the status of each labeled container is determined. This allows labeled containers that are unsatisfactory to be removed from the production process (*e.g.* information from the central database is communicated to robotic or human container handlers to remove the unsatisfactory sample). Likewise, containers that are automatically removed from the production process as unsatisfactory may be identified, and this information communicated to a central database (*e.g.* to update the status of an order, allow a re-order to be generated, etc). Allowing unsatisfactory samples to be removed prevents unnecessary manufacturing steps, and allows the production of a replacement to begin as early as possible.

As mentioned above, the tracking system of the present invention allows the production of single orders that have multiple components that may proceed through different production modules, and/or that may be processed (at least in part) in separate containers. For example, an order may be for the production of an INVADER detection kit. An INVADER detection kit is composed of at least 2 components (the INVADER oligonucleotide, and the downstream probe), and generally includes a second downstream probe (*e.g.* for a different allele), and one or two

synthetic targets so controls may be run (*i.e.* an INVADER kit may have 5 separate oligonucleotide sequences that need to be generated). The generation of separate sequences, in separate containers, generally necessitates that the tracking system track the location and status of each container, and direct the proper association of completed oligonucleotides into a single container or kit. Providing each container with a unique identifier corresponding to a single type of oligonucleotide (*e.g.* an INVADER oligonucleotide), and also corresponding to a single order (a SNP detection kit for diagnosing a certain SNP) allows separate, high through-put manufacture of the various components of a kit without confusion as to what components belong with each kit.

Tracking the location and status of the components of a kit (*e.g.* a kit composed of 5 different oligonucleotides) has many advantages. For example, near the end of the purification module HPLC is employed, and a simple sample analysis may be employed on each sample in each container to determine if a sample is collected in each tube. If no sample is collected after HPLC is performed, the unique identifier on the container, in connection with the central database, identifies the type of sample that should have been produced (*e.g.* INVADER oligonucleotide) and a re-order is generated. Identification of this particular oligonucleotide allows the manufacturing process for this oligonucleotide to start over from the beginning (*e.g.* this order gets priority status over other orders to begin the manufacturing process again). Importantly, the other components of the order may continue the manufacturing process without being discarded as part of a defective order (*e.g.* the manufacturing process may continue for these oligonucleotides up to the point where the defective oligonucleotide is required). Likewise, additional manufacturing resources are not wasted on the defective component (*i.e.* additional reagents and time are not spent on this portion of the order in further manufacturing steps).

The unique identifier on each of the containers allows the various components of a given order to be grouped together at a step when this is required (likewise, there is no need to group the components of an order in the manufacturing process until it is required). For example, prior to the dilute and fill module, the various components of a single order may be grouped together such that the contents of the proper containers are combined in the proper fashion in the dilute and fill module. This identification and grouping also allows re-orders to 'find' the other

components of a particular order. This type of grouping, for example, allows the automated mixing, in the dilute and fill stage, of the first and second downstream probes with the INVADER oligonucleotide, all from the same order. This helps prevent human errors in reading containers and accidentally providing probes intended for one SNP being labeled as specific for a different SNP (*i.e.* this helps prevent components of different kits from being accidentally mixed together). The identification of individual containers not only allows for the proper grouping of the various components of a single order, but also allows for an order to be customized for a particular customer (*e.g.* a certain concentration or buffer employed in the second dilute and fill procedure). Finally, containers with finished products in them (*e.g.* containers with probes, and containers with synthetic targets) need to be associated with each other so they are properly assayed in the quality control module, and packaged together as a single kit (otherwise, quality control and/or a final end-user may find false negative and false positives when attempting to test/use the kit). The ability to track the individual containers allows the components of a kit to be associated together by directing a robot or human operator what tubes belong together. Consequently, final kits are produced with the proper components. Therefore, the tracking systems and methods of the present invention allow high through-put production of kits with many components, while assuring quality production.

E. Inventory Control Component

In some embodiments, the present invention provides an inventory control component. In certain embodiments, the inventory control component comprises a computer system and one or more inventory components (e.g. cold storage facility, robotic assay component handling means, bar code scanners). In preferred embodiments, the computer system comprises enterprise application (e.g. ORACLE, PEOPLESOFT, BAAN, etc.) with a standard inventory control and material resource planning (MRP) software. In preferred embodiments, the inventory control system is configured to track and store (e.g. for weeks or months) detection assay components or full detection assays (e.g. all ready assembled into a kit). In some embodiments, the inventory control component handles (e.g. stores and retrieves when necessary) the detection assay components and detection assays by product number, or by product family, or by individual detection assay component.

In preferred embodiments, the inventory control component comprises a computer system operably linked to the other components (e.g. order entry components, detection assay centralized control network) such that inventory in the system can be tracked. This allows inventory to be displayed to a user placing an order, and allows the detection assay production component to be given real time instructions (e.g. a bill of material) to produce more detection assays (e.g. before inventory of particular assays or components becomes too low or falls to zero). Operably linking the inventory control component to the other systems of the present invention (see Data Management Systems in part IV below) allows raw materials to be ordered in a timely fashion facilitating effective supply chain management.

Also in preferred embodiments, the inventory control component comprises a cold storage area with coded (e.g. bar coded) detection assay components, and automated (e.g. robotic) storage and retrieval device. In some embodiments, the storage and retrieval device is configured to receive instructions (e.g. bill of material) from the computer system to store or retrieve various assay components, and assemble them into a desired detection assay. For example, the storage and retrieval device receive instructions to assemble the components of an INVADER assay. The device reads the codes on the various assay components stored in containers (e.g. on carousels) in the cold room to find the proper assay components (e.g. an INVADER oligonucleotide, a probe oligonucleotide, a FRET oligonucleotide, and a positive control target). In other embodiments, the components are stored and retrieved by location such that the containers do not need to be scanned (or they could be scanned to verify the correct assay component is selected). Once the storage and retrieval device obtains the desired components, they may be passed along to the Dilute and Fill component, or Packaging component for shipment to a customer.

F. Detection Assay Production Example

This Example describes the production of an INVADER assay kit for SNP detection using the automated DNA production system of the present invention.

1. Oligonucleotide Design

The sequence of the SNP to be detected is first submitted through the automated web-based user interface or through e-mail. The sequences are then transferred to the INVADER CREATOR software. The software designs the upstream INVADER oligonucleotide and downstream probe oligonucleotide. The sequences are returned to the user for inspection. At this point, the sequences are assigned a bar code and entered into the automated tracking system. The bar codes of the probe and INVADER oligonucleotide are linked so that their synthesis, analysis, and packaging can be coordinated.

2. Oligonucleotide Synthesis

Once the probe and INVADER oligonucleotide sequences have been designed, the sequences are transferred to the synthesis component. The bar codes are read and the sequences are logged into the synthesis module. Each module in this example consists of 14 MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), that prepare the primary probes, and two ABI 3900 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA), that prepare the INVADER oligonucleotides. Synthesizing a set of two primary and INVADER probes is complete 3-4 hours. The instruments run 24 h/day. Following synthesis, the automating tracking system reads the bar codes and logs the oligonucleotides as having completed the synthesis module.

The synthesis room is equipped with centralized reagent delivery. Acetonitrile is supplied to the synthesizers through stainless steel tubing. De-blocking solution (DCA in toluene) is supplied through Teflon tubing. Tubing is designed to attach to the synthesizers without any modification of the synthesizers. The synthesis room is also equipped with an automated waste removal system. Waste containers are equipped with ventilation and contain sensors that trigger removal of waste through centralized tubing when the cache pots are full. Waste is piped to a centralized storage facility equipped with a blow out wall. The pressure in the synthesis instruments is controlled with argon supplied through a centralized system. The argon delivery system includes local tanks supplied from a centralized storage tank.

During synthesis, the efficiency of each step of the reaction is monitored. If an oligonucleotide fails the synthesis process, it is re-synthesized. The bar coding system scans the container of the oligonucleotide and marks it as being sent back for re-synthesis.

Following synthesis, the oligonucleotides are transported to the cleavage and deprotection station. At this stage, completed oligonucleotides are subjected to a final deprotection step and are cleaved from the solid support used for synthesis. The cleavage and deprotection may be performed manually or through automated robotics. The oligonucleotides are cleaved from the solid support used for synthesis by incubation with concentrated NaOH and collected. The deprotection step takes 12 hours. Following cleavage and deprotection, the bar code scanner scans the oligonucleotide tubes and logs them as having completed the cleavage and deprotection step.

3. Purification

Following synthesis and cleavage, probe oligonucleotides are further purified using HPLC. INVADER oligonucleotides are not purified, but instead proceed directly to desalting (see below).

HPLC is performed on instruments integrated into banks (modules) of 8. Each HPLC module consists of a Leap Technologies 8-port injector connected to 8 automated Beckman-Coulter HPLC instruments. The automatic Leap injector can handle four 96-well plates of cleaved and deprotected primary probes at a time. The Leap injector automatically loads a sample onto each of the 8 HPLCs.

Buffers for HPLC purification are produced by the automated buffer preparation system. The buffer prep system is in a general access area. Prepared buffer is then piped through the wall in to clean room (HEPA environment). The system includes large vat carboys that receive premeasured reagents and water for centralized buffer preparation. The buffers are piped from central prep to HPLCs. The conductivity of the solution in the circulation loop is monitored as a means of verifying both correct content and adequate mixing. The circulation lines are fitted with venturis for static mixing of the solutions; additional mixing occurs as solutions are circulated through the piping loop. The circulation lines are fitted with 0.05 μm filters for sterilization and removal of any residual particulates.

Each purified probe is collected into a 50-ml conical tube in a carrying case in the fraction collector. Collection is based on a set method, which is triggered by an absorbance rate change within a predetermined time window. The HPLC is run at a flow rate of 5-7.5 ml/min (the maximum rate of the pumps is 10 ml/min.) and each column is automatically washed before the injector loads the next sample. The gradient used is described in Tables 3 and 4 and takes 34 minutes to complete (including wash steps to prepare the column for the next sample). When the fraction collector is full of eluted probes, the tubes are transferred manually to customized racks for concentration in a Genevac centrifugal evaporator. The Genevac racks, containing dry oligonucleotide, are then transferred to the TECAN Nap10 column handler for desalting.

4. Desalting

Following HPLC purification (probe oligonucleotides) or cleavage (INVADER oligonucleotides), oligonucleotides move to the desalting station. The dried oligonucleotides are resuspended in a small volume of water. Desalting steps are performed by a TECAN robot system. The racks used in Genevac centrifugation are also used in the desalting step, eliminating the need for transfer of tubes at this step. The racks are also designed to hold the different sizes of desalting columns, such as the NAP-5 and NAP-10 columns. The TECAN robot loads each oligonucleotide onto an individual NAP-5 or NAP-10 column, supplies the buffer, and collects the eluate.

5. Dilution

Following desalting, the oligonucleotides are transferred to the dilute and fill module for concentration normalization and dispenation. Each module consists of three automated probe dilution and normalization stations. Each station consists of a network-linked computer and a Biomek 2000 interfaced with a SPECTRAMAX spectrophotometer Model 190 or PLUS 384 (Molecular Devices Corp., Sunnyvale CA) in a HEPA-filtered environment.

The probe and INVADER oligonucleotides are transferred onto the Biomek 2000 deck and the sequence files are downloaded into the Biomek 2000. The Biomek 2000 automatically transfers a sample of each oligonucleotide to an optical plate, which the spectrophotometer reads to measure the A260 absorbance. Once the A260 has been determined, an Excel program

integrated with the Biomek software uses the measured absorbance and the sequence information to calculate the concentration of each oligonucleotide. The software then prepares a dilution table for each oligonucleotide. The probe and INVADER oligonucleotide are each diluted by the Biomek to a concentration appropriate for their intended use. The instrument then combines and dispenses the probe and INVADER oligonucleotides into 1.5 ml microtubes for each SNP set. The completed set of oligonucleotides contains enough material for 5,000 SNP assays.

If an oligonucleotide fails the dilution step, it is first re-diluted. If it again fails dilution, the oligonucleotide is re-purified or returned for re-synthesis. The progress of the oligonucleotide through the dilution module is tracked by the bar coding system. Oligonucleotides that pass the dilution module are scanned as having completed dilution and are moved to the next module.

6. Quality Control

Before shipping, the SNP set is subjected to a quality control assay in a SAGIAN CORE System (Beckman Coulter), which is read on a ABI 7700 real time fluorescence reader (PE Biosystems). The QC assay uses two no target blanks as negative controls and five untyped genomic samples as targets.

The quality control assay is performed in segments. In each segment, the operator or automated system performs the following steps: log on; select location; step specific activity; and log off. The ADS system is responsible for tracking tubes. If a tube is missing, existing ADS program routines will be used to discard/reorder/search for the tube.

In the first step, a picklist is generated. The list includes the identity of the SNPs that are being tested and the QC method chosen. The tubes containing the oligonucleotide are selected by the automated software and a copy of the picklist is printed. The tubes are removed from inventory by the operator and scanned with the bar code reader and being removed from inventory.

The operator or the automated system then takes the rack setup generated by the picklist and loads the rack. Tubes are scanned as they are placed onto the rack. The scan checks to make sure it is the correct tube and displays the location in the rack where the tube is to be placed. Completed racks are placed in a holding area to await the robot prep and robot run.

The operator or the automated system then chooses the genomics and reagent stock to be loaded onto the robot. The robot is programmed with the specific method for the SNP set generated. Lot numbers of the genomics and reagents are recorded. Racks are placed in the proper carousel location. After all the carousel locations have been loaded the robot is run.

Places are then incubated on the robot. The plates are placed onto heatblocks for a period of time specified in the method. The operator then takes the plate and loads it into the ABI 7700. A scan is started using the 7700 software. When the scan is completed the operator transfers the output file onto a Macintosh computer hard drive. The then starts the analysis application and scans in the plate bar code. The software instructs the operator to browse to the saved output file. The software then reads the file into the database and deletes the file.

The results of the QC assay are then analyzed. The operator scans plate in at workstation PC and reviews automated analysis. The automated actions are performed using a spreadsheet system. The automated spreadsheet program returns one of the following results:

1) Mark SNP Oligonucleotide ready for full fill (Operator discards diluted Probe/INVADER mixes. Requires no other action).
2) ReAssess Failed Oligonucleotide (Requires no action by operator, handled by automation).
3) Redilute Failed Oligonucleotide (Operator discards diluted tubes. Requires no other action).
4) Order Target Oligonucleotide (Requires no action by operator, handled by automation).
5) Fail Oligo(s) Discard Oligo(s) (Operator discards diluted tubes. Operator discards un-diluted tubes. Requires no other action).
6) Fail SNP (Operator discards diluted tubes. Operator discards un-diluted tubes. Requires no other action).
7) Full SNP Redesign (Operator discards diluted tubes. Operator discards un-diluted tubes. Requires no other action).
8) Partial SNP Redesign (Operator discards diluted tubes. Operator discards some un-diluted tubes. Requires no other action).

9) Manual Intervention (This step occurs if the operator or software has determined the SNP requires manual attention. This step puts the SNP "on hold" in the tracking system).

The operator then views each SNP analysis and either approves all automated actions, approves individual actions, marks actions as needing additional review, passes on reviewing anything, or over rides automated actions.

Once the SNP set has passed the QC analysis, the oligonucleotides are transferred to the packaging station.

In some embodiments, the produced detection assay is screened against a plurality of known sequences designed to represent one or more population groups, *e.g.*, to determine the ability of the detection assay to detect the intended target among the diverse alleles found in the general population. In preferred embodiments, the frequency of occurrence of the SNP allele in each of the one or more population groups is determined using the produced detection assay. Data collected may be used to satisfy regulatory requirements, if the detection assay is to be used as a clinical product.

IV. Data Management System

The present invention provides data management systems that integrate many of the components and systems of the present invention (See, e.g. Figures 58, 61 and 62). The data management systems of the present invention comprises networked computer processors (e.g. a local intranet), databases, and software applications that allow information to be shared and updated through the entire detection assay production and data collection process. The data management system may be comprised of the systems and components detailed above and below, all of which may be operably connected. This allows for integrated order entry, order analysis, assay design, assay production, inventory control, order shipping, and customer tracking, order tracking, inventory tracking, inventory control, and a product procurement module (e.g. that organizes ordering supplies from outside the company, or from within the same company, especially when manufacturing facilities are remote from one another). The data management systems of the present invention also facilitate other aspects of the present invention since information is constantly generated, evaluated, and stored (e.g. the rate of

development of ASRs and Clinical diagnostics is increased, See Product Development section below).

In yet another variant the system and method of the present invention provides a data feed that affects production of one or more oligonucleotide detection assays by the detection assay production component. Moreover, the detection assay production component, the shipping component, the shop floor control system, inventory control component and/or other components of the system can also receive the data feed from the web order entry component. In yet a further variant, the data feed may also be bi-directional or omni-directional between these various components of the system.

By way of example, the web order entry component data feed may provide input for routines that control and regulate the detection assay production component, the shipping component, the shop floor control system, inventory control component, other components of the system, and/or combinations thereof. In another aspect, there is a data feed from the detection assay production component, the shipping component, the shop floor control system, inventory control component, other components of the system, and/or combinations thereof to provide the consumer or other user information such as whether or not a detection assay is in stock, needs to be manufactured, lead times, shipping times, etc.

In other variants, the data feed comprises statistical information associated with one or more oligonucleotide detection assays. This statistical information can be created by various routines used by the system and methods from raw data obtained from the web order entry component, the detection assay production component, the shipping component, the shop floor control system, inventory control component, other components of the system, and/or combinations thereof. This information is then used in forecasting reagent supplies needed, and/or ordering other ingredients or components of the detection assays.

A generalized overview of certain embodiments of the data management systems of the present invention are provided in Figures 61 and 62. These figures show various computer systems, networks, and software applications of data management systems and how these components may be connected to facilitate the production of detection assays. These figures also show various components of the production facility, including certain production components, an inventory control system, and their relationship to order entry and processing components.

Figures 61 and 62 also demonstrate how the various computer systems, networks, and applications of the enterprise computer system are operably connected to the production components.

Referring specifically to Figures 61 and 62, initially an order is entered into the data management system by a client. This order may be a paper order (e.g. a contract for a large volume of assays), or it may be an electronic order placed through a web interface (e.g. INVADERCREATOR). Generally the order comprises a target sequence containing a SNP that a client wants to detect with a detection assay produced by the systems of the present invention. This sequence is entered into the system, which may come via a web order entry process when the data management system is operably linked to the world wide web. Preferably when oligonucleotides are ordered, a link to an accounting type database verifies that an active purchase order is in place to cover any assay development costs. Generally, a particular target is given a part number that is associated with the particular target to be detected. Then, as described below, an assay is designed for this target and tested, or multiple assays are designed and tested. Employing part numbers allows quick identification of which SNP is being detected (e.g. for future orders, and to quickly find where the SNP is located on a chromosome).

This target sequence is then analyzed. For example, this target sequence may already have a part number because it has previously been received by the systems of the present invention. In certain embodiments, this previously received target sequence skips target sequence analysis (e.g. in silico analysis, and assay design steps), and proceeds directly to job submit. In certain embodiments, target sequences that do not require analysis and assay design are marketed to clients at a reduced cost. Preferably, databases of the present invention have this information stored allowing newly entered sequences to be quickly searched. Also, the part number tracking of particular target SNPs allows information to be retrieved on how many assays have been designed for this target, and known confidence levels associated with each (which allows better and better assays to be developed for each target, and/or differential pricing for assays with different levels of confidence). For example, is a customer does not identify what SNP they are trying to identify, the assay design process will be run (potentially increasing the price of the assay) to validate the part number.

The part number validation process generally has three steps. First, once an order is received, the data management systems of the present invention determine if an assay has previously been designed for this SNP. Next, data is accessed (if available) that determines if the previously designed assay worked, and at what confidence level. Finally, a determination is made if there was ever a re-design of the assay, and if there is a master assay that has been designed (e.g. one that has been shown to work, and shown to work with an acceptable confidence level).

In circumstances where the sequence that is received does not match previously received target sequences (e.g. it is a custom order), the systems of the present invention may be configured to extensively analyze the target sequences for suitability. This process, known as in silico analysis involves three general steps. First, a preliminary screening step is performed that screens out repeat sequences, as well as artifacts such a vector sequences. Then, a database search is performed with the candidate target sequence to determine if the candidate sequence corresponds to a known sequence, contains a unique SNP to be detected, and that results from such detection are known to be reliable. Finally, this information if processed and/or stored. This information may be used to report the candidate target sequence as a "high probability sequence" (will allow the production of a valid detection assay), and this information provided to the client, or used to move the sequence along the data management system to a detection assay design step. Processing of this information may also reveal one or more problems with the candidate target sequence allowing a report to be sent (e.g. by the internet) to a user (*e.g.*, the person who input or requested the candidate target sequence or a technician utilizing the systems and methods of the present invention) highlighting the one or more problems.

If the target sequence is identified as a high probability sequence, or if the client requests that an assay be designed despite one or more problems, the target sequence information is forwarded (along the data management system) to the detection assay design systems of the present invention (e.g. comprising software applications to design assay components). In Figures 61 and 62, the detection assay design stage is represented with a long rectangular box containing "R-IC" (RNA INVADER CREATOR); "S-IC" (SNP INVADER CREATOR); "T-IC" (Transgene INVADER CREATOR); and "P-IC" (Primer INVADER CREATOR), as well as the design review box. Preferably, the data management system of the present invention has

software applications for designing the components of a detection assay. These software applications process the target sequence and generate appropriate designs for detection assays (e.g. INVADER assays, TaqMan Assays, multiplexed primers, etc.).

Figures 61 and 62 provide examples of software applications useful designing INVADER assays, and PCR primers for any type of detection assay. For example, S-IC (SNP INVADER CREATOR) is an example of software application that generates the preferred DNA probes (with appropriate flap), and INVADER oligonucleotides (See, A.II.B). Also, P-IC "Primer INVADER CREATOR) is an example of a software application able to generate highly multiplexed sets of PCR primers to be used in conjunction with other detection assays. Once appropriate designs are generated, these designs are moved (e.g. along the enterprise computer system) to the "job submit" stage. The job submit stage may be a database of assays that need to be fulfilled. As shown in Figures 61 and 62, these assays may already be in inventory, or may have to be produced (at least in part) by the production facility. Since the data management systems of the present invention integrate various components allows production and or inventory systems to be automatically activated (e.g. provided the correct instructions to begin assay production or to retrieve from storage, etc.).

If it is determined that the order can be filled from existing inventory, then many of the above steps may be skipped, and the order fulfilled from inventory. However, if it is determined that oligonucleotides need to be produced, the detection assay design is forwarded along the data management system (e.g. a work order or pick bill is generated) to the centralized control network that is operably connected to various production facility components (e.g. synthesis, cleave and deprotect) such that production is initiated.

Production may then begin with the oligonucleotide synthesis component. In preferred embodiments, more assays or components are generated than the work order actually requires (e.g. if one assay is ordered, ten are produced such that nine of the assays remain in inventory). In other preferred embodiments, the data management systems keep track of how many of each type of assays are produced and adjusts how many assays are made for inventory (e.g. keeping track of orders from individual customers or groups of customers allows forecasting of future orders, which may require that 20 assays are produced, instead of 10 assays, when inventory is depleted). In particular, instructions from the Centralized Control Network are sent to various

oligonucleotide synthesizers. The oligonucleotide synthesis component produces requested oligonucleotides, which are then transferred to the oligonucleotide processing components (e.g. cleavage and deprotection component, oligonucleotide purification, dilute and fill, quality control, and shipping or inventory control components; see figures 61 and 62). Preferably the tube, vials, and racks containing the requested oligonucleotides are labeled (e.g. with bar codes) such that the location of the oligonucleotides may be communicated to the centralized control network (and thus to other parts of the data management systems). This continues tracking allows all parts of the data management system to know in real time the status of particular orders. This information may be communicated back to the user (e.g. through a web interface, to customer service representatives, and to sales and business people), used to order raw materials, and used for business purposes.

Also information from the production facility, as shown in figures 61 and 62, may be communicated to the inventory control component. Preferably the inventory control component, as noted above, not only contains physical storage of previously manufactured oligonucleotides and assay (e.g. labeled with bar codes), but also comprises Enterprise Resource Planning (ERP) software having a standard MRP inventory control system. Any type of enterprise software may be employed (e.g. ORACLE, SAP, PEOPLESOFT, BAAN, etc.).

In certain embodiments, the data management system, when linked to the world wide web, provides additional information back to a user who is using the allele caller function. For example, an allele call may be made for a particular assay and this information provided to the user via the web. Also sent with the allele information may be links to information on public databases (e.g. papers on the clinical relevance of this particular SNP, unpublished clinical association studies, or links to internet pages describing certain drugs available for treatment of any disease associated with the SNP, or number of assays for this target remaining in inventory, or price discounts for this customer for re-order, other relevant products available, etc.). In certain embodiments, the information returned to the user associates a patient ID number with the allele call test result (e.g. sent via the web to a computer or a personal digital assistant). In preferred embodiments, the client ID number has medical history information associated with it such that allele calls help determine what SNPs are associated with a particular medical condition.

In certain embodiments, the data management system is operably linked to a customer's computer or computer system (e.g. via the world wide web). In this regard, the systems of the present invention may periodically (or continuously) query a customers computer system to determine if the customer requires additional detection assays to be shipped. For example, the data management system of the present invention may query a customer's computer (e.g. a database on the customer's computer or computer system) to determine if inventory is running low or is exhausted for any particular type of detection assay. Also, the customer's detection equipment may provide data to the customer's computer (e.g. the customer is running an allele caller on their computer). This data may also be queried by the systems of the present invention such that detection assays may be automatically ordered, or a prompt may be sent informing the customer of the availability of certain detection assays. For example, it the data generated by a customer that is stored on the customer's computer indicates that the customer will likely require certain panels of detection assays be designed, the systems of the present invention may communicate the availability of such assays (e.g. via email) to the customer. In this regard, the present invention provides a commercial advantage by allowing customer specific detection assays (and panels of assay) to be offered and/or sent to the customer in an automated fashion. This provides convenience and ease of use for the customer, and increased sales for supplies of assays. The detection assay may be any type of detection assay, including INVADER assays and TAQMAN assays. If additional assay are needed, the systems of the present invention may automatically design different/different assays for a customer, and suggestions for what the customer may want to order. For example, an email may be sent letting the customer know that their inventory is running low, or that their previously generated results will logically lead to further orders for additional assays. The system of the present invention may also design additional assays (e.g. TAQMAN or INVADER assays), or suggest alternative assays to the user (e.g. suggest an INVADER assay replace the TAQMAN assay previously employed by the user).

In preferred embodiments, the customer/user is part of the medical community (e.g. physician or lab using detection assays to provide results to physician). In some embodiments, the computer system is in a physician's office. A customer (e.g. physician) may have results of detection assay use sent to his or her computer (e.g. from the customer's detection equipment or from an outside lab). This information may be queried by the systems of the present invention,

which, as explained above, sends suggestions, alternative assays designs, or automatically sends detection assays. In further embodiments, information about what type of prescriptions a patient may require (e.g. based on the detection assay results) are provided to the physician (e.g. links to pages to order drugs that may required). In preferred embodiments, the detection assay reader device is located in the physician's office, and has a cost of less than ten thousand dollars. In preferred embodiments the patient's medical records are also used by the systems of the present invention to provide suggestions of prescriptions, and to suggest further detection assays that should be ordered (e.g. to avoid adverse drug reactions).

In certain embodiments, an electronic version of the Physicians Desk Reference (PDR), herein incorporated by reference, is available over the Internet. In preferred embodiments, the PDR may be queried by a user who is researching a particular condition. Preferably, the condition being queried by a user has information, or embedded information, that provides a user with particular detection assays that may be useful in diagnosing a disease, or confirming a disease, or to help avoid Adverse Drug Reactions with commonly prescribed medications. Preferably, the information regarding detection assays is operably linked to the Data Management Systems of the present invention. In this regard, one using the electronic PDR may be directed to an order screen to order the particular detection assays that may be required by the customer's patients.

V. Detection Assay Use and Data Generation and Collection

While the above sections describe the generation of a detection assay and the validation of the assay against a number of samples (e.g., several hundred samples), to fully investigate the viability of the detection assay against a broader population it is sometimes desired to conduct widespread testing with the detection assay. Where many different detection assays (e.g., hundreds to thousands of detection assays designed to identify unique markers) are to be investigated to facilitate moving products from research markets to clinical markets, large numbers of detection assays are tested against large numbers of samples.

In some embodiments, a detection assay producer distributes detection assays to research collaborators, whereby the research collaborators each conduct large numbers of tests (e.g., because of the inability of any one party to carry out a sufficient number of tests). The data

generated by these tests (e.g. returned to the data management system via the web) is used to validate the detection assay (e.g., for use in obtaining regulatory approval). Test results may show that the detection assay is suitable or not suitable for use in certain population sub-sets. The test results may also show that detection assays, for whatever reason (e.g., for determined or undetermined scientific reasons), are not suitable for one or more testing markets (e.g., do not provide the requisite data to achieve regulatory approval). Where tests are determined not suitable for a desired market, new tests may be generated using the methods described above to identify a candidate test that meets the desired criteria.

Information generated through use of detection assays may be collected and fed back into the data management system of the present invention. In this regard, ASRs and Clinical diagnostic products may be quickly identified. In some embodiments, the detection assays are shipped to a customer with an agreement that assay results will be reported back (e.g. thus reducing the price of the product, or automatically reported back through detection instruments linked to the world wide web).

In some embodiments, a detection assay directed to a single target is used. However, in certain preferred embodiments, panels containing a plurality of different detection assays are employed (e.g., produced and used in testing). For example, panels containing two or more markers associated with a particular medical condition are employed. In some preferred embodiments, the panels contain thousands of unique markers, corresponding to every identified medically relevant marker.

The present invention provides systems and methods to provide researchers using the detection assays with information to assist in data collection as well as system and methods to collect and analyze data. In particularly preferred embodiments, collected data is automatically directed to a processor for analysis, storage, and compilation (e.g., compilation to support an application requesting regulatory approval of clinical products).

In some such embodiments, the present invention provides users with a means to find known information (including but not limited to information gleaned from public sources, publications, patents, and information previously determined by any user of the database) about any SNP, other mutations, or other sequence characteristic that has been entered a database. In some embodiments, the present invention provides a facile means of linking known and collected

information about a particular SNP, other mutations, or other sequence characteristic to a particular test (*e.g.*, assay test) of a sample. The utility of such applications is illustrated below for embodiments where SNP information is to be analyzed.

A. Association databases

When a SNP has been linked to any other item of information (*e.g.*, disease state, chromosome location, gene, ethnic group, allele frequency, another SNP), it can be considered to have an association. Association databases may be configured with reference to any association or combination of associations. In a preferred embodiment, an association database is configured to contain information about SNPs that have been determined to have medical relevance (*i.e.*, to be relevant to some aspect of health, including but not limited to the presence of disease, disease susceptibility and prognosis, and individual response to particular therapy).

In one embodiment, information about a SNP can be provided in a database table (*e.g.*, a Microsoft Access database) having alphanumeric fields to provide details such as the gene identification, medical relevancy of the polymorphism, and literature or other references for the information provided (Figure 63). Any number of fields are contemplated. In some embodiments, information may be as simple as a single gene name or an accession number in a database (*e.g.*, GenBank). In other embodiments, the fields may provide more information, including but not limited to chromosome number, nucleotide, gene name, gene name abbreviation, genotype designation, allele location, GenBank accession number, NCBI URL link, dbSNP number, TSC number, targeted DNA sequence, disease category, disease association(s), SNP association(s) (*i.e.*, other SNPs or mutations found to be associated the SNP being reviewed), patent status (*e.g.*, whether a patent relating to that SNP has been identified), patent number(s), and the NCBI OMIM database URL link. Additional links or items of information may be provided, such as links to online reference libraries and patent or other intellectual property databases. Disease categories may include, for example, metabolism, endocrinology, pulminology, nephrology, gastroenterology, neurology, genetic disease, musculoskeletal, and immunology. Additional categories may be designated to specifically identify diseases that overlap into two or more particular categories. Yet another kind of category may be provided (*e.g.*, a "miscellaneous" category) for SNPs that have unknown or

indeterminate association, that have a known association that does not fall within another category, or that, for any other reason, are not appropriately assigned to another category. In some embodiments the database has one field. In preferred embodiments the database has at least 10 fields, and in a particularly preferred embodiment, the database has at least 20 fields. In some embodiments, the database table is displayed on a screen (Figure 63). In preferred embodiments, the screen is printable. In some embodiments, the fields are exportable to a spreadsheet file or worksheet (*e.g.*, in Microsoft Excel; Figure 64).

In one embodiment, the database may be searchable. In a preferred embodiment, the database is searchable, and is also configured to allow the user to present the resulting search data sets in an easily understandable, meaningful manner. In some embodiments, the database comprises an "allele caller" function, a function that provides allele calls (*i.e.*, identification of the alleles detected in a given assay) based on the data input (*e.g.*, such as from a fluorescent reader or mass spectrometer).

In some embodiments, the present invention provides a means for easily linking known information about a particular SNP to a particular test result on a sample through a "plate viewer" format corresponding to the layout of samples in a reaction vessel or plate (Figure 65). In preferred embodiments, the present information provides a means to use particular SNP test results on a sample to amend or update information about that SNP in an association database.

The following discussion provides one example of how a user interface for an association database may be configured. The user opens a work screen by clicking on an icon on a desktop display of a computer (*e.g.*, a Windows desktop). The work screen features a menu (*e.g.*, a drop down menu or "options" buttons) that allows the user to choose from available options. For example, in one embodiment, a user may be presented with the options of: 1) searching an association database; or 2) opening a plate viewer (as described above). In other embodiments, the user may have further or different options, such as 3) running an allele caller function. An option for exiting the program may be provided on the menu, as well. Examples of possible embodiments of user interfaces for each of these options are described, below.

1. Searching an association database:

In one embodiment, selecting this option opens a form having boxes that allow the user to make alphanumeric entries, and/or combination boxes (*e.g.*, boxes that allow the user to either select from a list or make an alphanumeric entry) for each field represented in that particular association database. The user can enter search criteria in any field or set of fields. Upon clicking a "search" button, the program constructs a query, searching for record sets that include the specified strings in the corresponding fields.

Matching records from the search are assembled into sets. In some embodiments, the matching sets are displayed on a screen. In other embodiments, the matching sets are exported (*e.g.*, sent to a printer or a file, or to a further process step) without display. In a preferred embodiment, the matching sets are displayed in a printable window.

In some embodiments, the user may select an entry from the matching set and view the information in the fields. In some embodiments, selection of an entry creates a display of the fields for that entry (Figure 66). In preferred embodiments, the fields are displayed in a new window. In other embodiments, the fields are exported (*e.g.*, sent to a printer or a file, or to a further process step) without display. In a preferred embodiment, the fields are displayed in a printable window. In some embodiments, one or more fields contain one or more local or Internet links (*e.g.*, hypertext links or URLs). In preferred embodiments, SNPs listed in a SNP association field provide links to the record(s) of the associated SNPs. In particularly preferred embodiments, the user can click on links to bring up the corresponding content.

2) Using a plate viewer

As noted above, the present invention provides a means for easily linking known information about a particular SNP to a particular test result on a sample through a "plate viewer" format, *i.e.*, in a fashion that corresponds to (*e.g.*, visually represents) the layout of samples in a reaction vessel (Figure 65). For example, if test assays for SNPs are performed in 96-well microtiter plates, which are arranged in grids of 8 wells X 12 wells, the links to the information regarding the SNPs would be displayed in a grid of 8 X 12 cells, such that each cell corresponds to the particular well in the plate (*i.e.*, the test SNP in the 3rd well of the 4th row will have a link to its information presented on screen in the 3rd cell of the 4th row). Similar displays

corresponding to other layouts of reaction vessels are contemplated (*e.g.*, staggered grids, or circular or linear layouts). Any layout that can be replicated as a computer display is contemplated, including any non-gridded, or random distribution of reaction vessels in any arrangement that may be captured for representation on a computer display. Locations may be entered manually, or they may be automatically sensed and entered by methods such as digital imaging, coordinate sensing (*e.g.*, such as that used for touch-screen computer displays), and the like.

Using a 384-well plate, a user selecting a "Plate Viewer" option should be presented with a table in the 384-well plate layout. In one embodiment, the SNPs entered into each cell of the table are assigned by the user (*e.g.*, by entering identifying information from a particular field, such as a dbSNP number, into a selected cell on the plate viewer table). In preferred embodiments, SNPs are pre-assigned to particular cells. In particularly preferred embodiments, the SNPs are pre-assigned to cells in the table such that they correspond with an assay plate configured to test those SNPs in the corresponding wells. In other particularly preferred embodiments, the user selects from a menu of Plate Viewers, each having a different set of SNPs in pre-assigned cells corresponding with an assay plate configured to test those SNPs in the corresponding wells.

In one embodiment, the user selects which field of the SNP record assigned to that cell will be displayed in the cell. In some embodiments, different fields from each SNP record may be displayed in each of the different cells. In other embodiments, the cells are coordinated so that the same field from each SNP record is displayed in each assigned cell. In a preferred embodiment, the user can globally change the fields displayed in all cells (*e.g.*, through the use of a menu), such that all of the cells can be changed at one time to display the same field from each different SNP record.

In some embodiments, there is a code to visually distinguish test SNPs from control reactions (*e.g.*, 'no target' controls or other controls). In preferred embodiments, the code is a color code.

In some embodiments, the user may select an entry from a cell and view (*e.g.*, in a "data viewer") the information in all of the fields for that SNP record (Figure 66). In some embodiments, selection of an entry creates a display of the fields for that entry. In preferred

embodiments, the fields are displayed in a new window. In other embodiments, the fields are exported (*e.g.*, sent to a printer or a file, or to a further process step) without display. In a preferred embodiment, the fields are displayed in a printable window. In some embodiments, one or more fields contain one or more local or Internet links (*e.g.*, hypertext links or URLs). In preferred embodiments, the user can click on links to bring up the corresponding content.

In some embodiments, an association database is provided on removable storage media (*e.g.*, compact disc). In further embodiments, the storage media having the database includes an index of any PlateViewers having pre-assigned SNP records contained thereon. In preferred embodiments, the storage media having the database provides an indication of the currency of the information in the recorded database (*e.g.*, a date or date range, version number, etc.). In preferred embodiments, the storage media having the database provides contact information for technical support (*e.g.*, phone numbers facsimile numbers, email addresses, street addresses, names of technical support personnel, etc.).

B). Running an allele caller function.

In some embodiments, the association database comprises an "allele caller" function, a function that provides identification of the alleles detected in a given assay, based on input assay data (*e.g.*, from an instrument such as a fluorescent reader, nucleic acid chip reader, or mass spectrometer).

The data to be processed by an allele caller may be provided in many different forms. In some embodiments, the data is raw signal, such as number corresponding to a measurement of fluorescence signal from a spot on a chip or a reaction vessel, or a number corresponding to measurement of a peak (*e.g.*, peak height or area, as from, for example, a mass spectrometer, HPLC or capillary separation device). In some embodiments the data is imported directly from a measuring device. In other embodiments, the data is imported from a file. Raw data may be generated by any number of SNP detection methods, including but not limited to those listed below.

1. Direct sequencing Assays

In some embodiments of the present invention, variant sequences are detected using a direct sequencing technique. In these assays, DNA samples are first isolated from a subject using any suitable method. In some embodiments, the region of interest is cloned into a suitable vector and amplified by growth in a host cell (*e.g.*, a bacteria). In other embodiments, DNA in the region of interest is amplified using PCR.

Following amplification, DNA in the region of interest (*e.g.*, the region containing the SNP or mutation of interest) is sequenced using any suitable method, including but not limited to manual sequencing using radioactive marker nucleotides, or automated sequencing. The results of the sequencing are displayed using any suitable method. The sequence is examined and the presence or absence of a given SNP or mutation is determined.

2. PCR Assay

In some embodiments of the present invention, variant sequences are detected using a PCR-based assay. In some embodiments, the PCR assay comprises the use of oligonucleotide primers that hybridize only to the variant or wild type allele (*e.g.*, to the region of polymorphism or mutation). Both sets of primers are used to amplify a sample of DNA. If only the mutant primers result in a PCR product, then the patient has the mutant allele. If only the wild-type primers result in a PCR product, then the patient has the wild type allele.

3. Fragment Length Polymorphism Assays

In some embodiments of the present invention, variant sequences are detected using a fragment length polymorphism assay. In a fragment length polymorphism assay, a unique DNA banding pattern based on cleaving the DNA at a series of positions is generated using an enzyme (*e.g.*, a restriction enzyme or a CLEAVASE I [Third Wave Technologies, Madison, WI] enzyme). DNA fragments from a sample containing a SNP or a mutation will have a different banding pattern than wild type.

a. RFLP Assay

In some embodiments of the present invention, variant sequences are detected using a restriction fragment length polymorphism assay (RFLP). The region of interest is first isolated using PCR. The PCR products are then cleaved with restriction enzymes known to give a unique length fragment for a given polymorphism. The restriction-enzyme digested PCR products are generally separated by gel electrophoresis and may be visualized by ethidium bromide staining. The length of the fragments is compared to molecular weight markers and fragments generated from wild-type and mutant controls.

b. CFLP Assay

In other embodiments, variant sequences are detected using a CLEAVASE fragment length polymorphism assay (CFLP; Third Wave Technologies, Madison, WI; *See e.g.*, U.S. Patent Nos. 5,843,654; 5,843,669; 5,719,208; and 5,888,780; each of which is herein incorporated by reference). This assay is based on the observation that when single strands of DNA fold on themselves, they assume higher order structures that are highly individual to the precise sequence of the DNA molecule. These secondary structures involve partially duplexed regions of DNA such that single stranded regions are juxtaposed with double stranded DNA hairpins. The CLEAVASE I enzyme, is a structure-specific, thermostable nuclease that recognizes and cleaves the junctions between these single-stranded and double-stranded regions.

The region of interest is first isolated, for example, using PCR. In preferred embodiments, one or both strands are labeled. Then, DNA strands are separated by heating. Next, the reactions are cooled to allow intrastrand secondary structure to form. The PCR products are then treated with the CLEAVASE I enzyme to generate a series of fragments that are unique to a given SNP or mutation. The CLEAVASE enzyme treated PCR products are separated and detected (*e.g.*, by denaturing gel electrophoresis) and visualized (*e.g.*, by autoradiography, fluorescence imaging or staining). The length of the fragments is compared to molecular weight markers and fragments generated from wild-type and mutant controls.

4. Hybridization Assays

In preferred embodiments of the present invention, variant sequences are detected a hybridization assay. In a hybridization assay, the presence of absence of a given SNP or mutation is determined based on the ability of the DNA from the sample to hybridize to a complementary DNA molecule (*e.g.*, a oligonucleotide probe). A variety of hybridization assays using a variety of technologies for hybridization and detection are available. A description of a selection of assays is provided below.

a. Direct Detection of Hybridization

In some embodiments, hybridization of a probe to the sequence of interest (*e.g.*, a SNP or mutation) is detected directly by visualizing a bound probe (*e.g.*, a Northern or Southern assay; *See e.g.*, Ausabel *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, NY [1991]). In a these assays, genomic DNA (Southern) or RNA (Northern) is isolated from a subject. The DNA or RNA is then cleaved with a series of restriction enzymes that cleave infrequently in the genome and not near any of the markers being assayed. The DNA or RNA is then separated (*e.g.*, on an agarose gel) and transferred to a membrane. A labeled (*e.g.*, by incorporating a radionucleotide) probe or probes specific for the SNP or mutation being detected is allowed to contact the membrane under a condition or low, medium, or high stringency conditions. Unbound probe is removed and the presence of binding is detected by visualizing the labeled probe.

b. Detection of Hybridization Using "DNA Chip" Assays

In some embodiments of the present invention, variant sequences are detected using a DNA chip hybridization assay. In this assay, a series of oligonucleotide probes are affixed to a solid support. The oligonucleotide probes are designed to be unique to a given SNP or mutation. The DNA sample of interest is contacted with the DNA "chip" and hybridization is detected.

In some embodiments, the DNA chip assay is a GeneChip (Affymetrix, Santa Clara, CA; *See e.g.*, U.S. Patent Nos. 6,045,996; 5,925,525; and 5,858,659; each of which is herein incorporated by reference) assay. The GeneChip technology uses miniaturized, high-density arrays of oligonucleotide probes affixed to a "chip." Probe arrays are manufactured by

Affymetrix's light-directed chemical synthesis process, which combines solid-phase chemical synthesis with photolithographic fabrication techniques employed in the semiconductor industry. Using a series of photolithographic masks to define chip exposure sites, followed by specific chemical synthesis steps, the process constructs high-density arrays of oligonucleotides, with each probe in a predefined position in the array. Multiple probe arrays are synthesized simultaneously on a large glass wafer. The wafers are then diced, and individual probe arrays are packaged in injection-molded plastic cartridges, which protect them from the environment and serve as chambers for hybridization.

The nucleic acid to be analyzed is isolated, amplified by PCR, and labeled with a fluorescent reporter group. The labeled DNA is then incubated with the array using a fluidics station. The array is then inserted into the scanner, where patterns of hybridization are detected. The hybridization data are collected as light emitted from the fluorescent reporter groups already incorporated into the target, which is bound to the probe array. Probes that perfectly match the target generally produce stronger signals than those that have mismatches. Since the sequence and position of each probe on the array are known, by complementarity, the identity of the target nucleic acid applied to the probe array can be determined.

In other embodiments, a DNA microchip containing electronically captured probes (Nanogen, San Diego, CA) is utilized (*See e.g.*, U.S. Patent Nos. 6,017,696; 6,068,818; and 6,051,380; each of which are herein incorporated by reference). Through the use of microelectronics, Nanogen's technology enables the active movement and concentration of charged molecules to and from designated test sites on its semiconductor microchip. DNA capture probes unique to a given SNP or mutation are electronically placed at, or "addressed" to, specific sites on the microchip. Since DNA has a strong negative charge, it can be electronically moved to an area of positive charge.

First, a test site or a row of test sites on the microchip is electronically activated with a positive charge. Next, a solution containing the DNA probes is introduced onto the microchip. The negatively charged probes rapidly move to the positively charged sites, where they concentrate and are chemically bound to a site on the microchip. The microchip is then washed and another solution of distinct DNA probes is added until the array of specifically bound DNA probes is complete.

A test sample is then analyzed for the presence of target DNA molecules by determining which of the DNA capture probes hybridize, with complementary DNA in the test sample (*e.g.*, a PCR amplified gene of interest). An electronic charge is also used to move and concentrate target molecules to one or more test sites on the microchip. The electronic concentration of sample DNA at each test site promotes rapid hybridization of sample DNA with complementary capture probes (hybridization may occur in minutes). To remove any unbound or nonspecifically bound DNA from each site, the polarity or charge of the site is reversed to negative, thereby forcing any unbound or nonspecifically bound DNA back into solution away from the capture probes. A laser-based fluorescence scanner is used to detect binding,

In still further embodiments, an array technology based upon the segregation of fluids on a flat surface (chip) by differences in surface tension (ProtoGene, Palo Alto, CA) is utilized (*See e.g.*, U.S. Patent Nos. 6,001,311; 5,985,551; and 5,474,796; each of which is herein incorporated by reference). Protogene's technology is based on the fact that fluids can be segregated on a flat surface by differences in surface tension that have been imparted by chemical coatings. Once so segregated, oligonucleotide probes are synthesized directly on the chip by ink-jet printing of reagents. The array with its reaction sites defined by surface tension is mounted on a X/Y translation stage under a set of four piezoelectric nozzles, one for each of the four standard DNA bases. The translation stage moves along each of the rows of the array and the appropriate reagent is delivered to each of the reaction site. For example, the A amidite is delivered only to the sites where amidite A is to be coupled during that synthesis step and so on. Common reagents and washes are delivered by flooding the entire surface and then removing them by spinning.

DNA probes unique for the SNP or mutation of interest are affixed to the chip using Protogene's technology. The chip is then contacted with the PCR-amplified genes of interest. Following hybridization, unbound DNA is removed and hybridization is detected using any suitable method (*e.g.*, by fluorescence de-quenching of an incorporated fluorescent group).

In yet other embodiments, a "bead array" is used for the detection of polymorphisms (Illumina, San Diego, CA; *See e.g.*, PCT Publications WO 99/67641 and WO 00/39587, each of which is herein incorporated by reference). Illumina uses a BEAD ARRAY technology that combines fiber optic bundles and beads that self-assemble into an array. Each fiber optic bundle

contains thousands to millions of individual fibers depending on the diameter of the bundle. The beads are coated with an oligonucleotide specific for the detection of a given SNP or mutation. Batches of beads are combined to form a pool specific to the array. To perform an assay, the BEAD ARRAY is contacted with a prepared subject sample (e.g., DNA). Hybridization is detected using any suitable method.

c. Enzymatic Detection of Hybridization

In some embodiments of the present invention, hybridization is detected by enzymatic cleavage of specific structures (INVADER assay, Third Wave Technologies; *See e.g.*, U.S. Patent Nos. 5,846,717, 6,090,543; 6,001,567; 5,985,557; and 5,994,069; each of which is herein incorporated by reference). The INVADER assay detects specific DNA and RNA sequences by using structure-specific enzymes to cleave a complex formed by the hybridization of overlapping oligonucleotide probes. Elevated temperature and an excess of one of the probes enable multiple probes to be cleaved for each target sequence present without temperature cycling. These cleaved probes then direct cleavage of a second labeled probe. The secondary probe oligonucleotide can be 5'-end labeled with a fluorescent dye that is quenched by a second dye or other quenching moiety. Upon cleavage, the de-quenched dye-labeled product may be detected using a standard fluorescence plate reader, or an instrument configured to collect fluorescence data during the course of the reaction (*i.e.*, a "real-time" fluorescence detector, such as an ABI 7700 Sequence Detection System, Applied Biosystems, Foster City, CA).

The INVADER assay detects specific mutations and SNPs in unamplified genomic DNA. In an embodiment of the INVADER assay used for detecting SNPs in genomic DNA, two oligonucleotides (a primary probe specific either for a SNP/mutation or wild type sequence, and an INVADER oligonucleotide) hybridize in tandem to the genomic DNA to form an overlapping structure. A structure-specific nuclease enzyme recognizes this overlapping structure and cleaves the primary probe. In a secondary reaction, cleaved primary probe combines with a fluorescence-labeled secondary probe to create another overlapping structure that is cleaved by the enzyme. The initial and secondary reactions can run concurrently in the same vessel. Cleavage of the secondary probe is detected by using a fluorescence detector, as described above. The signal of the test sample may be compared to known positive and negative controls.

In some embodiments, hybridization of a bound probe is detected using a TaqMan assay (PE Biosystems, Foster City, CA; *See e.g.*, U.S. Patent Nos. 5,962,233 and 5,538,848, each of which is herein incorporated by reference). The assay is performed during a PCR reaction. The TaqMan assay exploits the 5'-3' exonuclease activity of DNA polymerases such as AMPLITAQ DNA polymerase. A probe, specific for a given allele or mutation, is included in the PCR reaction. The probe consists of an oligonucleotide with a 5'-reporter dye (*e.g.*, a fluorescent dye) and a 3'-quencher dye. During PCR, if the probe is bound to its target, the 5'-3' nucleolytic activity of the AMPLITAQ polymerase cleaves the probe between the reporter and the quencher dye. The separation of the reporter dye from the quencher dye results in an increase of fluorescence. The signal accumulates with each cycle of PCR and can be monitored with a fluorimeter.

In still further embodiments, polymorphisms are detected using the SNP-IT primer extension assay (Orchid Biosciences, Princeton, NJ; *See e.g.*, U.S. Patent Nos. 5,952,174 and 5,919,626, each of which is herein incorporated by reference). In this assay, SNPs are identified by using a specially synthesized DNA primer and a DNA polymerase to selectively extend the DNA chain by one base at the suspected SNP location. DNA in the region of interest is amplified and denatured. Polymerase reactions are then performed using miniaturized systems called microfluidics. Detection is accomplished by adding a label to the nucleotide suspected of being at the SNP or mutation location. Incorporation of the label into the DNA can be detected by any suitable method (*e.g.*, if the nucleotide contains a biotin label, detection is via a fluorescently labelled antibody specific for biotin).

5. Other Detection Assays

Additional detection assays that are produced and utilized using the systems and methods of the present invention include, but are not limited to, enzyme mismatch cleavage methods (e.g., Variagenics, U.S. Pat. Nos. 6,110,684, 5,958,692, 5,851,770, herein incorporated by reference in their entireties); polymerase chain reaction; branched hybridization methods (e.g., Chiron, U.S. Pat. Nos. 5,849,481, 5,710,264, 5,124,246, and 5,624,802, herein incorporated by reference in their entireties); rolling circle replication (e.g., U.S. Pat. Nos. 6,210,884 and 6,183,960, herein incorporated by reference in their entireties); NASBA (e.g., U.S. Pat. No. 5,409,818, herein

incorporated by reference in its entirety); molecular beacon technology (e.g., U.S. Pat. No. 6,150,097, herein incorporated by reference in its entirety); E-sensor technology (Motorola, U.S. Pat. Nos. 6,248,229, 6,221,583, 6,013,170, and 6,063,573, herein incorporated by reference in their entireties); cycling probe technology (e.g., U.S. Pat. Nos. 5,403,711, 5,011,769, and 5,660,988, herein incorporated by reference in their entireties); Dade Behring signal amplification methods (e.g., U.S. Pat. Nos. 6,121,001, 6,110,677, 5,914,230, 5,882,867, and 5,792,614, herein incorporated by reference in their entireties); ligase chain reaction (Barnay Proc. Natl. Acad. Sci USA 88, 189-93 (1991)); and sandwich hybridization methods (e.g., U.S. Pat. No. 5,288,609, herein incorporated by reference in its entirety).

6. Mass Spectroscopy Assay

In some embodiments, a MassARRAY system (Sequenom, San Diego, CA.) is used to detect variant sequences (*See e.g.*, U.S. Patent Nos. 6,043,031; 5,777,324; and 5,605,798; each of which is herein incorporated by reference). DNA is isolated from blood samples using standard procedures. Next, specific DNA regions containing the mutation or SNP of interest, about 200 base pairs in length, are amplified by PCR. The amplified fragments are then attached by one strand to a solid surface and the non-immobilized strands are removed by standard denaturation and washing. The remaining immobilized single strand then serves as a template for automated enzymatic reactions that produce genotype specific diagnostic products.

Very small quantities of the enzymatic products, typically five to ten nanoliters, are then transferred to a SpectroCHIP array for subsequent automated analysis with the SpectroREADER mass spectrometer. Each spot is preloaded with light absorbing crystals that form a matrix with the dispensed diagnostic product. The MassARRAY system uses MALDI-TOF (Matrix Assisted Laser Desorption Ionization - Time of Flight) mass spectrometry. In a process known as desorption, the matrix is hit with a pulse from a laser beam. Energy from the laser beam is transferred to the matrix and it is vaporized resulting in a small amount of the diagnostic product being expelled into a flight tube. As the diagnostic product is charged when an electrical field pulse is subsequently applied to the tube they are launched down the flight tube towards a detector. The time between application of the electrical field pulse and collision of the diagnostic product with the detector is referred to as the time of flight. This is a very precise

measure of the product's molecular weight, as a molecule's mass correlates directly with time of flight with smaller molecules flying faster than larger molecules. The entire assay is completed in less than one thousandth of a second, enabling samples to be analyzed in a total of 3-5 second including repetitive data collection. The SpectroTYPER software then calculates, records, compares and reports the genotypes at the rate of three seconds per sample.

In some embodiments, data generated by different detection methods are processed to facilitate comparison, *e.g.*, using an process like the Extraction-Transformation-Load paradigm from Data Warehousing, wherein data is "published" into a single repository, normalizing disparate data, and optimizing it for browsing and easy access to normalized, integrated data (*e.g.*, DataMart and MetaSymphony software, NetGenics, Inc., Cleveland OH; US Patent 6,125,383, incorporated herein by reference in its entirety). SNP data generated by one SNP analysis method may be compared to SNP results data generated by another SNP analysis method (*e.g.*, INVADER assay results are compared to gene chip data).

In some embodiments of the present invention, data is processed using an algorithm selected to determine an allele from the input assay data. The algorithm selected for processing data may be determined by the nature of the input assay data. The following provides an example of the application of an allele caller to an assay run in a microtiter plate (*e.g.*, a 384-well plate).

The user enters information to identify the plate to be analyzed. In one embodiment, the plate may be identified by entry of a code number (*e.g.*, a barcode number, part number, lot number). In another embodiment, the program provides a menu from which the user selects the number corresponding to the plate.

In some embodiments, the program provides a validation of the plate. For example, in some embodiments, the program verifies that the plate is of a suitable format for available analysis (*e.g.*, that it corresponds to an assay for which an allele caller function can be provided). In other embodiments, the program verifies that the plate has been passed through some other process step. In some embodiments wherein the association database is provided on removable media (*e.g.*, as described above), the program verifies that the version of the CD in use is suitable (*e.g.*, has an appropriate version of an allele caller function, or has an appropriate association database) for use with the plate to be analyzed.

When a plate has been identified and determined to be valid for analysis, a record is displayed. In preferred embodiments, the record is a table having cells that correspond to assay wells on a microtiter plate (*e.g.*, a "plate viewer", described above). In some embodiments, the user has the option (*e.g.*, through a menu selection) of creating a new analysis record or of calling up a record of a prior analysis. In preferred embodiments, the record links to identifying data from other analyses performed on the same collection of samples (*e.g.*, name, date generated, etc.). In particularly preferred embodiments, SNP test wells on a plate are linked through a "plate viewer" function to SNP records in a database. In further particularly preferred embodiments, the database is an association database.

Prior to analysis, the assay data from the plate is imported, or "loaded" into the analysis program. It is contemplated that the data to be processed by an allele caller may be provided in many different forms. In some embodiments, the assay data is raw (*i.e.*, unanalyzed) signal, such as a number corresponding to a measurement of fluorescence signal from a spot on a chip or a reaction vessel, or a number corresponding to measurement of a peak (*e.g.*, peak height or area, as from, for example, a mass spectrometer, HPLC or capillary separation device). In some embodiments the data is imported directly from a measuring device. In other embodiments, the data is imported from a file. Raw assay data may be generated by any number of SNP detection methods, including but not limited to those listed above.

In some embodiments, the loaded assay data is displayed on a screen. In preferred embodiments, data is displayed in a plate viewer format. In some preferred embodiments, the layout is displayed in a new window. In particularly preferred embodiments, the window is printable.

Loaded assay data is then analyzed or processed using one or more algorithms selected to determine an allele from the input assay data. The algorithm selected for processing data is generally determined by the nature of the input assay data. In some embodiments, analysis involves determining the presence or absence of a signal (*e.g.*, detectable fluorescence, or a detectable peak). In other embodiments, analysis involves determining the presence of a signal meeting a threshold value. In still other embodiments, analysis involves a comparison of more than one signal (*e.g.*, examining differences in signal level, calculating ratios, etc.). In preferred embodiments, a SNP result (*i.e.*, a determination of genotype at that locus, such as homozygous

Allele 1 or Allele 2, heterozygous, Indeterminate) is determined when the processed data yields or corresponds to a value that has been predetermined to be indicative of a particular SNP result.

In some embodiments, the SNP results data from one plate are compared with the SNP results data from another plate. In other embodiments, SNP results data generated by one SNP analysis source method are compared to SNP results data generated by another SNP analysis method (*e.g.*, INVADER assay results are compared to gene chip data).

In some embodiments, analysis results are displayed. In other embodiments, the analysis results are exported (*e.g.*, sent to a printer or a file, or to a further process step) without display. In preferred embodiments, SNP results are displayed on a screen. In particularly preferred embodiments, results are displayed in a plate viewer (Figures 67 and 68). In some preferred embodiments, the plate viewer is displayed in a new window. In particularly preferred embodiments, the window is printable.

In some embodiments, the user may select a particular SNP result from the display of results and view the information in fields. In some embodiments, selection of an entry creates a display of the fields for that entry. In some embodiments, all the fields of the SNP record in an association database are shown. In other embodiments, a subset of the fields is shown. In preferred embodiments, fields in SNP results records include but are not limited to results of the analysis (*e.g.*, homozygous Allele 1 or Allele 2, heterozygous, Indeterminate), the entered or imported raw input assay data (*e.g.*, measured fluorescence, measured peaks, etc.), or the analyzed input assay data by which the allele determination was made (*e.g.*, calculated differences in signal level, calculated ratios). In preferred embodiments, a field for user comments is included. In particularly preferred embodiments, the user comment field is editable after a SNP result has been obtained. In further particularly preferred embodiments, changes in a SNP result record may be saved by the user to that record or to a version of that record after a comment field is edited.

In some embodiments, the user selects which field of the SNP result record assigned to that cell will be displayed in the cell (Figures 67 and 68). In some embodiments, different fields from each SNP result record may be displayed in each of the different cells. In other embodiments, the cells are coordinated so that the same field from each SNP result record is displayed in each assigned cell. In a preferred embodiment, the user can globally change the

fields displayed in all wells (*e.g.*, through the use of a menu), such that all of the cells can be changed at one time to display the same field from each different SNP result record.

In preferred embodiments, the fields are displayed in a new window. In other embodiments, the fields are exported (*e.g.*, sent to a printer or a file, or to a further process step) without display. In a preferred embodiment, the fields are displayed in a printable window. In some embodiments, one or more fields will contain one ore more local or Internet links (*e.g.*, hypertext links or URLs). In preferred embodiments, the user can click on links to bring up the corresponding content.

In some embodiments, there is a code to visually distinguish test SNPs results and control reaction results (*e.g.*, 'no target' controls or other controls). In preferred embodiments, the code is a color code.

In some embodiments, the fields are exportable to a spreadsheet file or worksheet (*e.g.*, in Microsoft Excel, Figure 69). In some embodiments, SNP result data are exported to a worksheet by field content (*e.g.*, one worksheet with all allele calls, one worksheet with all calculated ratios of signals, one worksheet with all raw input fluorescence measurements). In other embodiments, SNP results data are exported, all data is exported to a single worksheet, with data grouped according to the well with which it corresponds. In preferred embodiments, the user has the option (*e.g.*, through a menu or window) of selecting a variety ways in which the SNP results data are sorted and/or grouped for export to a spreadsheet.

In preferred embodiments, following verification, assays for the detection of a given SNP are tested on a plurality of additional individuals. Data from additional assays is combined with information obtained from database searches. In preferred embodiments, the result is a revised reliability score for the SNP. In particularly preferred embodiments, data from additional analysis (*e.g.*, results generated by an investigator using the methods and systems of the present invention) is used to update or amend an association database containing information about the given SNP.

C. Database Software

In some embodiments, GENOMICA (Boulder, CO) software is utilized to generate and host the SNP database of the present invention, which may be located, for example, on the data

management systems of the present invention. In some embodiments, GENOMICA DISCOVERY MANAGER software is utilized. Genomica software utililizes Oracle databases to provide a web interface, security features, and reporting information (e.g., including but not limited to, the information described in Section C below). Depending on the particular application, one or more of the features of DISCOVERY MANAGER are utilized.

D. Revisions of Database Information

In preferred embodiments, the information (e.g., reliability scores) in the SNP database of the present invention is revised on a regular basis. In some embodiments, the revisions are automated. For example, users (e.g., customers) provide data from genotyping studies (e.g., through an automated web interface). In some embodiments, individual users are given a reliability rating based on the quality of their genotyping information. In preferred embodiments, the contribution to the reliability score of an individual's data is weighted based on the reliability rating of the user. In addition, individual databases are given reliability ratings based on the verification of their data.

E. Automated Genotyping

In preferred embodiments, the detection assays are employed in an automated or semi-automated fashion (e.g. a detection assay readout requires minimal human interaction), such that high throughput genotyping may be achieved. Any type of automated genotyping system of platform may be employed. In preferred embodiments, the automated genotyping systems of the present invention comprise at least one liquid handling platform, at least one detection platform, and at least one incubation component. Table 2 provides examples of such genotyping systems useful with the present invention.

TABLE 2

System	Liquid Handler	Detection	Incubation	Robotics
CyBio	CyBi-well 384s (3)	TECAN Saffire	Liconic StoreX 200 or Heraeus 6070	convey or rail
Packard Plate Track	384 MPD (3)	TECAN SAFFIRE	Liconic StoreX 200 or Heraeus 6070	convey or rail
Beckman CORE w/FX	Biomek F/X (2 Arm-384)	Perspective Cytoflurs 4000 or LJL Analyst	Liconic StoreX 200 or Heraeus 6070	OCRA 3M rail
Packard Minitrack	384 MPD (1)	TECAN Saffire	Liconic StoreX 200 or Heraeus 6070	convey or rail
CyBio	CyBi-Well 384s (2)	TECAN Saffire	Liconic StoreX 200 or Heraeus 6070	convey or rail
TECAN workstation	TECAN Genesis 200 +/- M'mek (96)	TECAN Spectrofluor +	Liconic StoreX 44/200	ROMA
Beckman CORE w/BK2	Biomek 200 +/- M'mek (96)	Perspective Cytoflurs 4000 or LJL Analyst	Liconic StoreX 44/ 200 or Heraeus 6070	ORCA 3M rail

Other types of automated equipment and systems may be used with the systems of the present invention to facilitate high throughput genotyping. Other useful systems include Robbins, Cartesian, and Zymar systems. Exemplary liquid handling platforms include, but are not limited to; Beckman Coulter Biomek 200, Beckman Coulter Biomek FX, Beckman Coulter Multimek, CyBio CyBiWell 384, CyBio CyBiDrop, TECAN Genesis, 100, 150, 200 platforms, Cartesian Technologies SynQuad Systems, Zymark Sciclone ALH, Robbins Tango 384, Packard Multiprobe I and II, and Packard Mini & Plate Trak systems. Examplary detection platforms include, but are not limited to, Bio-Tek FL800, Perseptive Cytofluor 4000, Tecan Genios, Tecan Spectrafluor Plus, PE Wallac Victor, BMG Fluorostar, Packard Fusion, Tecan Saffire, Tecan Ultra, LJL Analyst, and Packard Image Trak. Examplary Incubation components include, but are not limited to, manual incubation components including, but not limited to, Heat Blocks (e.g. 96 well plate), Thermalcyclers (e.g. used in incubator), Bio-Ovens (e.g. 10 plate), and Heraeus UT 6060 (e.g. 30 plate). Exemplary incubation components that are automation friendly include,

but are not limited to, Liconic Store X 40 (e.g. 44 plate), Heraeus Cytomat 2 (e.g. 42 plate), Liconic StoreX 200 (e.g. 200 plate), and Heraeus Cytomat 6070 (e.g.189 plate).

An example of a protocol for set up of 96 and/or 384-well INVADER assays using the BIOMEK 2000 CORE system is shown in Figure 59A. Also, Figures 59B, 59B, and 59C also show exemplary automated genotyping systems useful for high throughput screening. Further exemplary configurations for automated genotyping systems include, but are not limited to, the following five configurations: 1) System: Beckman Sagain CORE system, Robotics: Beckman Sagian 3m ORCA, Liquid Handler: Beckman Biomek 2000, Plate Washer Biomek 2000 WASH-8 tool, Incubation (75C):Dry Bath Heat Blocks, Incubation (60C): Heraeus Cytomat 6070 Automated Incubator, Reader: Perseptive Cytofluor 4000; 2) System: Beckman Sagian CORE system, Robotics: Beckman Sagian 3M ORCA, Liquid Handler: Beckman Biomek FX, Dual bridge with 96 and Span-8 channel pipettor heads, Plate Washer: Bio-Tek, Molecular Devices, etc., Incubation (75C): Liconic StoreX44 or Heraeus CytoMat2 Automated incubators, Incubator (60C): Liconic StoreX44 or Heraeus CytoMat2 Automated incubator, Reader: TECAN Safire, Spectrafluor, Ultra, or the like; 3) Robotics: Beckman Sagian, 2M Orca robot, Liquid Handler: Beckman Biomek FX, Dual bridge system with Span-8 and 384 pipette heads, Incubator: Heraeus Cytomat 6070, Reader: Tecan Safire Monochromator, Plate Storage: Beckman ambient carousel; 4) Robotics: Beckman Sagian, Coneyor Alps and onboard Gripper, Liquid Handler: Beckman FX, Dual bridge system with Span-8 and 384 pipette heads, Incubator: Heraeus Cytomat 6070, Reader; Tecan Safire Monochromator, Plate Storage: Heraeus Cytomat hotel (ambient); and 5) Robotics: Integral plate conveyors and rotating transfer arms, Liquid Handler: (3) CyBi Well 384 pipettors, and (1) CyBiDrop pipettor, Incubator: Liconic StoreX200, Reader: Tecan Safire Monochromator, and Plate Storage: CyBio high capacity plate stackers. Preferably, the automated genotyping systems of the present invention have a capacity of 50-75,000 genotypes per day in 384 well plates. In other preferred embodiments, the automated genotyping systems of the present invention have a capcity of at least 150,000, or at least 200,000 (e.g. aproximately 200,000) per day. It is understood that the automated genotying systems may require some off line plate arraying of either sample or probes to allow 384-channel pipetting and plate transfers to occur on the high throughput line.

F. Determination of Allele Frequencies in Pooled Samples

In particular embodiments, the present invention allows detection of polymorphims in pooled samples combined from many individuals in a population (e.g. 10, 50, 100, or 500 individuals), or from a single subject where the nucleic acid sequences are from a large number of cells that are assayed at once. In this regard, the present invention allows the frequency of rare mutations in pooled samples to be detected and an allele frequency for the population established. In some embodiments, this allele frequency may then be used to statistically analyze the results of applying the INVADER detection assay to an individual's frequency for the polymorphism (e.g. determined using the INVADER assay). In this regard, mutations that rely on a percent of mutants found (e.g. loss of heterozygozity mutations) may be analyzed, and the severity of disease or progression of a disease determined (See, e.g. US Patent 6,146,828 and 6,203,993 to Lapidus, hereby incorporated by reference for all purposes, where genetic testing and statistical analysis are employed to find disease causing mutations or identify a patient sample as containing a disease causing mutations).

In some embodiments of the present invention, broad population screens are performed. In some preferred embodiments, pooling DNA from several hundred or a thousand individuals is optimal. In such a pool, for example, DNA from any one individual would not be detectable, and any detectable signal would provide a measure of frequency of the detected allele in a broader population. The amount of DNA to be used, for example, would be set not by the number of individuals in a pool, as was done in the 15-person pool described in Example 3, but rather by the allele frequency to be detected. For example, the assay in the 96-well format would give ample signal from 20 to 40 ng of DNA in a 90 minute reaction. At this level of sensitivity, analysis of 1 μg of DNA from a high-complexity pool would produce comparable signal from alleles present in only about 3-5% of the population. In some embodiments, reactions are configured to run in smaller volumes, such that less DNA is required for each analysis. In some preferred embodiments, reactions are performed in microwell plates (*e.g.*, 384-well assay plates), and at least two alleles or loci are detected in each reaction well. In particularly preferred embodiments, the signals measured from each of said two or more alleles or loci in each well are compared.

Pooled Sample - Example 1

This example describes the detection of a polymorphism in the APOC4 gene. In particular, this example describes the use of the INVADER assay to detect a mutation in the APOC4 gene in pooled samples.

In this example, genomic DNAs were isolated from blood samples from several individual donors, and were characterized by invasive cleavage for the T/C polymorphism in codon 96 of the *APOC4* gene (*See*, Allan, *et al., Genomics* 1995 Jul 20;28(2):291-300, hereby incorporated by reference). The *APOC4* assay used 5' GATTCGAGGAACCAGGCCTTGGTGT (SEQ ID NO:1) 3' as the invasive oligonucleotide and either 5' ATGACGTGGCAGACAGCGGACCCAGGTCC-$PO_4$3' (SEQ ID NO:2) or 5' ATGACGTGGCAGACCGCGGACCCAGGTCC-$PO_4$3' (SEQ ID NO:3) as primary signal probes for the T (Leu96) and the C (Pro96) alleles, respectively. The secondary target and probe were 5' CGGAGGAAGCGTTAGTCTGCCAC<u>GTCAT</u>-NH_2 3' (SEQ ID NO:4) and 5' FAM-TAAC[Cy3]GCTTCCTGCCG 3', respectively (SEQ ID NO:5).

All oligonucleotides were synthesized using standard phosphoramidite chemistries. Primary probe oligonucleotides were unlabeled. The FRET probes were labeled by the incorporation of Cy3 phosphoramidite and fluorescein phosphoramidite (Glen Research, Sterling, VA). While designed for 5' terminal use, the Cy3 phosphoramidite has an additional monomethoxy trityl (MMT) group on the dye that can be removed to allow further synthetic chain extension, resulting in an internal label with the dye bridging a gap in the sugar-phosphate backbone of the oligonucleotide. Amine or phosphate modifications, as indicated, were used on the 3' ends of the primary probes and the secondary target oligonucleotides to prevent their use as invasive oligonucleotides. 2'-O-methyl bases in the secondary target oligonucleotides are indicated by underlining and were also used to minimize enzyme recognition of 3' ends. Approximate probe melting temperatures (T_ms) were calculated using the Oligo 5.0 software (National Biosciences, Plymouth, MN); non-complementary regions were excluded from the calculations.

Pooled samples were constructed by diluting the heterozygous (het) DNA into DNA that is homozygous T (L96) at this locus. The test reactions contained 0.08 to 8 μg of T (L96) genomic DNA per reaction, and the het DNA was held at 0.08 μg, thus creating a set of mixtures

in which het DNA represented from 50% down to 1% of the total DNA in the sample (See, Figure 70). The actual representation of the C (P96) allele ranged from 25% down to 0.5% of the copies of this gene in the mixed samples. Controls included reactions having either all T (L96) DNA at each of the various DNA levels, or all het DNA at the 80 ng level. In addition, a sample of DNA that is homozygous for the C (P96) allele was tested (Fig. 2).

For all the INVADER assay reactions, 4 pmol of invasive probe, 40 pmol of FRET probe, and 20 pmol of secondary target oligonucleotide were combined with genomic DNA in 34 μl of 10 mM MOPS (pH 7.5) with 1.6 % PEG. Reactions with the C (Pro96) allele of the *APOC4* gene contained 80 ng of DNA heterozygous for this allele, and included DNA homozygous for the T (Leu96) allele at the indicated ratios. Samples were overlaid with 15 μl of Chill-Out liquid wax and heated to 95°C for 5 min to denature the DNA. Upon cooling to 67°C the reactions were started by the addition of 400 ng of Cleavase VIII enzyme, 15 pmol of either the T (Leu96) or the C (Pro96) primary signal probe, and $MgCl_2$ to a final concentration of 7.5 mM. The plates were incubated for 2 hours at 67°C, cooled to 54°C to initiate the secondary (FRET) reaction, and incubated for another 2 hours. The reactions were then stopped by addition of 60 μl of TE. The fluorescence signals were measured on a Cytofluor fluorescence plate reader at excitation 485/20, emission 530/25, gain 65, temperature 25° C. Three replicates were done for each reaction and for no-target controls. The average signal for each target DNA was calculated, the average background from the no-target controls was subtracted, and the data plotted using Microsoft Excel.

The results of this example are shown in Figure 70. As shown in this figure, the C (P96) allele was easily detected in all reactions, including that in which it was present in only 0.5% of the *APOC4* alleles present in the mixture. These data indicate that the invasive cleavage reactions can be used for population analysis using pooled DNA samples. This has the double advantage of reducing the number of assays required to verify a new SNP, and of allowing the use of one large preparation of pooled DNA for numerous tests, thereby reducing the influence of sample-to-sample variations in DNA purity.

The above example demonstrates that the INVADER assay may be used to screen a population. A sample of mixed DNA to be analyzed should be large enough to bring the low-frequency alleles into the detectable range, *e.g.*, 80 to 100 ng of the variant genome in these 40

μl reactions. As shown above in this Example, a sample of 8 to 10 μg of mixed DNA allowed detection of alleles present at 0.5 to 1% of the population under these conditions. In addition, the DNA from any one individual ideally should not be present in a large enough quantity to generate a detectable signal when an aliquot of the pool is tested. Creating a pool of several hundred individuals should guarantee that any detected signal reflects a contribution from many individuals in the pool. Finally, the use of a second probe set as an internal standard would allow the signals to be normalized from reaction to reaction, and would allow the prevalence of any SNP to be measured more accurately.

Pooled Sample - Example 2

This example describes the detection of a polymorphism in the CFTR gene. In particular, this example describes the use of the INVADER assay to detect the ΔF508 mutation in the CFTR gene in a pooled sample.

For INVADER assay analysis of the ΔF508 mutation, the primary probe set comprised 5' ATATTCATAGGAAACACCAAG 3' (SEQ ID NO:6) as the invasive oligonucleotide and either 5' AACGAGGCGCACAGATGATATTTTCTTTAA 3' (SEQ ID NO:7) or 5' ATCGTCCGCCTCTGATATTTTCTTTAATGG 3' (SEQ ID NO:8) as signal probes for the wild type and the mutant alleles. The secondary reaction components were designed to function optimally at a temperature at least 5 degrees below the primary reaction temperature.

All oligonucleotides described were synthesized using standard phosphoramidite chemistries. Primary probe oligonucleotides were unlabeled. The FRET probes were labeled by the incorporation of Cy3 phosphoramidite and fluorescein phosphoramidite (Glén Research, Sterling, VA). While designed for 5' terminal use, the Cy3 phosphoramidite has an additional monomethoxy trityl (MMT) group on the dye that can be removed to allow further synthetic chain extension, resulting in an internal label with the dye bridging a gap in the sugar-phosphate backbone of the oligonucleotide. One nucleotide was omitted at this position to accommodate the dye. Amine modifications were used on the 3' ends of the primary probes, the secondary target and the arrestor oligonucleotides to prevent their use as invasive oligonucleotides. 2'-O-methyl bases are indicated by underlining and are also used to minimize enzyme recognition of 3' ends. Approximate probe melting temperatures were calculated using the Oligo 5.0 software

(National Biosciences, Plymouth, MN); noncomplementary regions were excluded from the calculations.

DNA samples characterized for CFTR genotype were purchased from Coriell Institute for Medical Research (Camden, NJ), catalog numbers NA07469 (heterozygous in the CFTR gene for both ΔF508 and R553X mutations) and NA01531 (homozygous ΔF508). To determine what dose of a mutant could be detected within a pooled sample using the FRET-sequential invasive cleavage approach, DNA that is the heterozygous for the ΔF508 mutation in the CFTR gene was diluted into DNA that is homozygous wild type at that locus. The test reactions contained 0.1to 2.6 μg of the total genomic DNA per reaction, and the mutant DNA was held at 0.1 μg, thus creating a set of mixtures in which mutant DNA represented from 50% down to 4% of the total DNA in the sample. Because the mutant DNA was heterozygous at the 508 locus, the actual allelic representation ranged from 25% down to 2% of the DNA in the mixed samples. Controls included reactions having either all wt at each of the various DNA levels, or all heterozygous mutant DNA at the 100 ng level. In addition, a sample of DNA that is homozygous for the ΔF508 mutation was tested.

DNA concentrations were estimated using the PicoGreen method. 4 pmol of INVADER probe, 40 pmol of FRET probe, and 20 pmole of secondary target oligonucleotide were combined with genomic DNA in 34 μl of 10 mM MOPS (pH 7.5) with 4% PEG. Samples were overlaid with 15 μl of Chill-Out liquid wax and heated to 95°C for 5 min to denature the DNA. Upon cooling to 62°C the reactions were started by the addition of 400 ng of AfuFEN1 enzyme, 15 pmole of either wt or mutant primary probe, and $MgCl_2$ to a final concentration of 7.5 mM. The plates were incubated for 2 hours at 62°C, cooled to 54°C to initiate the secondary (FRET) reaction, and incubated for another 2 hours. The reactions were then stopped by addition of 60 μl of TE. The fluorescence signals were measured on a Cytofluor fluorescence plate reader excitation 485/20, emission 530/25, gain 65, temperature 25° C. Three replicates were done for each reaction and for no-target controls. The average signal for each target DNA was calculated, the average background from the no-target controls was subtracted, and the data plotted using Microsoft Excel.

The results of this Example are presented in Figure 71. Analysis of the signal from the mutant allele shows that it is not noticeably inhibited by substantial increases in the amount of

wild type DNA, and the ΔF508 mutant DNA could be easily detected when present as only 2% of the mixture (Figure 71). These data indicate that the invasive cleavage reactions can be used for population analysis using pooled DNA samples. This has the double benefit of reducing the number of assays required to verify a new SNP, and of allowing the use of one large, preparation of the pooled DNA to be used for numerous tests, thereby reducing the influence of sample-to-sample variations in DNA purity.

Application of the INVADER assay to screen populations is possible given the results presented in this example. In preferred embodiments for population screening, the DNA contribution from each individual should be equal, and the DNA from any one individual should not be present in a large enough quantity to generate a detectable signal when an aliquot of the pool is tested. For example, for this system creating a large enough pool that any one person contributes less than 1 ng (e.g., 0.5 ng) to each reaction should guarantee that any detected signal reflects a contribution from many individuals in the pool. For other detection systems, limiting the DNA from any one individual to an amount less than the detection limit of the system, for example 1/5 to 1/10 the detection limit, should produce the desired effect. The use of a second probe set as an internal standard, for example, would allow the signals to be normalized from reaction to reaction, and would allow the prevalence of any SNP to be measured more accurately.

Pooled Sample - Example 3

This example describes the detection of the Consortium No. TSC 0006429 (SNP 1831) mutation in pooled samples. DNA from 15 individuals was purchased from the Coriell Cell Repository and each sample was tested to identify the genotype at the SNP Consortium No. TSC 0006429 (SNP 1831) locus. Each reaction contained 40 ng of DNA from each individual, 0.366 μM primary probe. 0.0366 μM Invader oligonucleotide, 0.183 μM FRET Probe and 100 ng CLEAVASE VIII enzyme in a buffer of 10 mM MOPS (pH 7.5) with 7.5 mM $MgCl_2$.

The probes used were as follows (5' to 3'):

Invader: CTTACTTGACCTTGGGCCCAGTTATTTAACCTTCTAGACCT (SEQ ID NO:9); Probe T: CGCGCCGAGGATCAGTTTCTTCATCTCTAAAATGGA (SEQ ID NO:10); Probe G: CGCGCCGAGGCTCAGTTTCTTCATCTCTAAAATGGA (SEQ ID NO:11); Synthetic Target T: TGTATCCATTTTAGAGATGAAGAAACTGAG (SEQ ID NO:12);

GGTCTAGAAGGTTAAATAACTGGGCCCAAGGTCAAGTAAGGG (SEQ ID NO:13);
Synthetic Target G: TGTATCCATTTTAGAGATGAAGAAACTGAT (SEQ ID NO:14);
GGTCTAGAAGGTTAAATAACTGGGCCCAAGGTCAAGTAAGGG (SEQ ID NO:15)

The assays were performed as described in Hall *et al., PNAS*, 97 (15):8272 (2000). Briefly, reaction were incubated at a constant temperature of 65° C. The data for each sample, produced using an ABI 7700 instrument for real-time reaction detection, are shown in the 15 panels of Figures 72 and 73, with signals from the G allele shown as the light line and from the T allele shown as the dark line. The signal from each allele present in the mixture appears as an ascending curve reflecting the quadratic nature of the signal accumulation; the signal from any allele not present is essentially a straight line. These DNAs were then pooled in several combinations: Samples 1-5, 6-10, 11-15, 1-10, 6-15, and 1-15. The data panels are shown in Figure 74. Figure 75 provides a comparison of the net fluorescence counts measured at the end of each reaction. From the results in 66a-b, the allele representation in each mixture can be calculated. Both Figures 74 and 75 demonstrate that the aggregate signals for each pool are proportional with respect to the final ratio of the alleles in the mix. The net fluorescence signals from the pooled samples are greater than those from the individuals because the amount of DNA from each person was held constant. For example, the assays run on DNA pooled from 5 individuals had 5 times as much DNA as the assays run on DNA from one individual.

As seen in this example, the real-time detection capabilities of the ABI 7700 can prove invaluable in detecting rare SNPs. Because the reaction is a two-step cascade, the real-time trace of signal accumulated in the Invader assay fits to a quadratic equation (*i.e.*, the curves observed in Figures 72, 73, and 74), but background signal remains linear over the course of the reaction. Consequently, distinguishing signal arising from the genomic target from the background fluorescence is straightforward. This characteristic of the assay means that low-level signals from rare alleles can be resolved from background with more certainty.

Pooled Sample - Example 4

Measurement of different alleles within a single reaction removes concerns about sample-to-sample variations introducing inaccuracies into the measurements to be compared in the determination of allele frequency. Use of biplex (detection of two alleles or loci per reaction) or

more complex multiplex (detection of more than two alleles or loci per reaction) configurations increases the through-put for allele frequency determination and facilitates comparisons of allele frequencies between different populations (*e.g.*, affected vs. non-affected with a particular trait).

The following provides one example of a general protocol for the detection of two alleles in a DNA sample, and several examples wherein the protocol has been applied to the determination of alleles in samples. In this example, the signals are measured from fluorescein dye (FAM) and REDMOND RED dye (Red, Synthetic Genetics, San Diego, CA), each used on a separate FRET probe in combination with the Z28 ECLIPSE quencher (Synthetic Genetics, San Diego, CA). This protocol is provided to serve as an example and is not intended to limit the use of the methods or compositions of the present invention to any particular assay protocol or reaction configuration. Numerous fluorescent dyes and fluorophore/quencher combinations, and the methods of attaching and detecting such agents alone and in FRET combinations to nucleic acids are known in the art. Such other agents combinations are contemplated for use in the present invention and their use in these methods is within the scope of the present invention.

a. Procedure for Allele Frequency Determination in Pooled DNA

1. Determine the DNA concentration of each of the samples to be used in the INVADER Assay using the PICOGREEN reagents (procedure follows).
2. Mix the DNA samples at the desired ratios to mimic pools of genomic samples at specified allelic frequencies.
3. Denature the genomic DNA samples by incubating them at 95° C for 10 min. Sample may then be placed on ice (optional).
4. Prepare a Probe/INVADER oligonucleotide /$MgCl_2$ mix by combining the 1.15 µL probe/INVADER oligonucleotide mix (3.5 µM of each primary probe and 0.35 µM INVADER oligonucleotide) and the 1.85 µL 24 mM $MgCl_2$ per reaction. Preparation of a master mix sufficient for testing of the complete set of samples is preferred.
5. Add 3µl of the appropriate control or sample DNA target at 80 to 100 ng/µl (approximately 240-300 ng of genomic DNA) to the appropriate well of a 384-well biplex INVADER Assay FRET detection plate (Third Wave Technologies, Madison, WI). Each plate well contains 3

μl of a solution, dried after dispensing, containing 10 mM MOPS, 8% PEG, 4% glycerol, 0.06 % NP 40, 0.06% Tween 20, 12 ug/ml BSA, 50 ng/ul BSA, 33.3 ng/ul CLEAVASE VIII enzyme, 1.17 μM FAM FRET probe (5' – FAM-TCT (Z28) AG CCG GTT TTC CGG CTG AGA GTC TGC CAC GTC AT - 3', SEQ ID NO:16) and 1.17μM Red FRET Probe (5' – Red-TCT (Z28) TC GGC CTT TTG GCC GAG AGA CCT CGG CGC G - 3', SEQ ID NO:17).

6. Next, pipette 3 μl of Probe/INVADER oligonucleotide/$MgCl_2$ mix into the appropriate wells of the 384-well biplex INVADER Assay FRET detection plate.

7. Overlay each reaction with 6 μL of mineral oil.

8. Cover the plates with an adhesive cover and spin at 1,000rpm in a Beckman GS-15R centrifuge (or equivalent) for 10 seconds to force the probe and target into the bottom of the wells.

9. Incubate the reactions at 63°C for 3-4 hours in a thermal cycler or incubator such as a BioOven III. After 3-4h incubation at 63°C, lower the temperature to 4°C if a thermalcycler is being used or to RT if an incubator is being used.

10. Analyze the microtiter plate on a fluorescence plate reader using the following parameters:

		Wavelength/Bandwidth
FAM:	Excitation:	485nm/20nm
	Emission:	530nm/25nm

		Wavelength/Bandwidth
Red:	Excitation:	560nm/20nm
	Emission:	620nm/40nm

b. Calculation of fold-over-zero minus 1 (FOZ-1):

The signals from each reaction are measured by comparison to the signal from a no-target control (the 'zero') and are expressed as a multiple of the signal from the 'zero' reaction. The

factor one is subtracted to get the factor of actual signal over the background (*e.g.*, for a sample having 1.5 X the signal of the zero or 1.5 fold-over-zero, the amount of specific signal is 1.5-1, or 0.5).

Determine FOZ-1 as follows:

<u>FOZ-1 FAM Probe</u> = ((raw counts FAM probe 1, 485/530) / (raw counts from No Target Control FAM probe , 485/530)) –1 .

<u>FOZ-1 Red Probe</u> = ((raw counts Red probe 2, 560/620) / (raw counts from No Target Control Red probe, 560/620)) –1

c. Calculation the Correction Factor (CF) as follows

A correction factor can be calculated to accommodate any variations in the efficiencies of the cleavage reactions between the probe sets.

$\mathbf{CF_{FAM}} = (FOZ_{FAM}\text{-}1) / (FOZ_{Red}\text{-}1)$; $\mathbf{CF_{Red}} = (FOZ_{Red}\text{-}1) / (FOZ_{FAM}\text{-}1)$ of a heterozygous control.

For the FAM allelic frequency calculation:

$$\frac{(FOZ_{FAM}-1) / CF_{FAM}}{((FOZ_{FAM}-1) / CF_{FAM}) + (FOZ_{Red} - 1)} \quad x100$$

For the Red allelic frequency calculation:

$$\frac{(FOZ_{Red}-1) / CF_{Red}}{((FOZ_{Red}-1) / CF_{Red}) + (FOZ_{FAM} - 1)} \quad x100$$

d. DNA quantitation procedure (Molecular Probes PICOGREEN Assay)

The PICOGREEN reagent is an asymmetrical cyanine dye (Molecular Probes, Eugene, OR). Free dye does not fluoresce, but upon binding to dsDNA it exhibits a >1000-fold

fluorescence enhancement. PICOGREEN is 10,000-fold more sensitive than UV absorbance methods, and highly selective for dsDNA over ssDNA and RNA.

1. Turn on the fluorescence plate reader at least 10 minutes before reading results. Use the following settings to read the PICOGREEN results:

	Wavelength/Bandwidth
Excitation	~485nm/20nm
Emission:	~530nm/25nm

2. Prepare 1X TE buffer (10mM Tris-HCl, 1mM EDTA, pH 7.5) from the 20X TE stock which is supplied in the PICOGREEN kit (to make 50ml, add 2.5ml of 20X TE to 47.5ml sterile, distilled DNase-free water). 50ml is sufficient for 250 assays.
3. Dilute DNA standards from 100μg/ml to 2μg/ml with 1X TE. For two standard curves, prepare 400μl of a 2μg/ml stock by adding 8μl of the 100μg/ml stock to 392μl 1X TE.
4. Prepare the two standard curves in the microtiter plate as shown in the table 7:

TABLE 7

Plate Well	Final [DNA] (ng/ml)	Vol. (μl) 2μg/ml DNA Standard	Vol. (μl) 1X TE Buffer
A1 & A2	0	0	100
B1 & B2	25	2.5	97.5
C1 & C2	50	5	95
D1 & D2	100	10	90
E1 & E2	200	20	80
F1 & F2	300	30	70
G1 & G2	400	40	60

H1 & H2	500	50	50

5. For each unknown, add 2µl of sample to 98µl of 1X TE in the microplate well. Mix by pipetting up and down.
6. Prepare a 1:200 dilution of the PICOGREEN reagent in 1X TE. For each standard and each unknown sample, a volume of 100µl is needed. For example, 2 standard curves with 8 points each will require 1.6 ml. To calculate the total volume of diluted PICOGREEN reagent needed, determine the total number of samples and unknowns will be tested and multiply this number by 100µl (if using a multichannel pipet, make extra reagent). The PICOGREEN reagent is light sensitive and should be kept wrapped in foil while thawing and in the diluted state. Vortex well.
7. Add 100µl of diluted PICOGREEN to every standard and sample. Mix by pipetting up and down.
8. Cover the microplate with foil and incubate at room temperature for 2–5 minutes.
9. Read the plate.
10. Generate a standard curve using the average values of the standards and determine the concentration of DNA in the unknown samples.

e. Measurement of allele frequencies in genomic DNA samples

DNA samples having alleles at various frequencies were created by mixing different homozygous genomic DNA samples at different ratios. Each pool contained a total of 240 ng genomic DNA, and the reactions were carried out in 384-well plates as described above, at 63°C for 3 hours. The measured signals are shown in Figure 76A. The allelic frequencies were calculated based on the relative signal generated by the FAM and Red reporter dyes, and are displayed graphically in Figure 76B. These data show the correlation between the theoretical or actual allelic frequency (the frequency intended to be created by mixing known amounts of DNA), compared to the allelic frequency calculated from the INVADER assay data.

An 8-way pool of the genomic DNA of different individual was also tested. Each of the 8 DNA was previously characterized for each of 8 different SNP loci, so that the allelic frequency for each of the 8 SNPs in the pool was known. In this test, each pool contained a total

of 300 ng genomic DNA, and the reactions were carried out in 384-well plates as described above, at 63°C for 3 hours. The measured signals for the FAM channel, the rarer allele in each case, is shown in Figure 77. The graph compares the known frequencies for each allele to the frequencies calculated from the INVADER assay data.

DNAs homozygous for each of two different SNPs (SNP132505 and SNP131534) were combined at various ratios to simulate genomic pools with different allelic frequencies. Each pool contained a total of 240 ng genomic DNA, and the reactions were carried out in 384-well plates as described above, at 63°C for 3 hours. The allelic frequencies were calculated based on the relative signal generated by the FAM and Red reporter dyes, and are displayed graphically in Figures 78A and 78B.

The probes used in the tests described above and additional probes sets suitable for use in the methods of the invention are shown in Figure 80A-C.

VI. Integrated Information, Design, and Production

Data gathered from the use of detection assays on one or more samples (e.g., as described in Section V, above) may be used to generate and expand powerful genomics databases and to supplement and improve target selections, detection assay design, detection assay productions, and detection assay use, and further analysis of detection assay results. The data may also be used to obtain regulatory approval for clinical products for detection assays that are demonstrated to meet the necessary requirements for clinical regulatory approval (described below). While, for clarity, each of the components of the systems and methods of the present invention have generally been described herein in isolation, each component relates to each other component, and the synergy between the components provides enhanced systems and methods for acquiring and analyzing biological information. This synergy, as it relates to some embodiments of the present invention, is represented in Figure 81. The center of the figure shows genomic databases representing phenotypic databases (e.g., disease databases), genomic databases (e.g., genome sequence databases, polymorphism databases, allele frequency databases, etc.), and expressed RNA databases. Data in the databases is derived from any number of sources. For example, the databases may contain data from compiled public or private databases. Data may also be actively incorporated using systems and methods of the

present invention. As shown in Figure 81, data is received from investigators (e.g., using a communication network) providing target sequence requests for in silico analysis, detection assay design, and/or detection assay production (See e.g., Sections AI, AII, and AIII, above).

In some embodiments, new data is generated during the processes of the present invention (*e.g*, produced assays may be tested on a plurality of samples to determine allele frequencies, as described in Section AIII). New data is also received from detection assay data gathered from investigators (See *e.g.*, Section AV, above). In some embodiments of the present invention, information is tracked and correlated from the initial target sequence requests to the final detection assay result data analysis.

Newly collected data may be incorporated into a number of aspects of the present invention. It can be used to refine in silico analysis, *e.g.*, to provide improved output information; it may be added to an association database, *e.g.*, to note newly observed associations within existing fields, and/or to define new fields indicating new types of associations, such as allele frequency within populations tested.

The following example is provided to illustrate certain preferred embodiments of the present invention. In this example, the systems for performing in silico analysis, detection assay design and production, and information management and analysis are provided by a service provider. Target sequences to be analyzed are provided by a first user (e.g., a researcher, pharmaceutical company, government agency, etc.) and detection assays generated to detect the target sequence are used by the first user and/or other users.

The first user selects a target sequence of interest. For example, an investigator may have identified a SNP in a human genomic sequence that is correlated to disease state (e.g., a SNP correlated to cardiovascular disease, diabetes, development of cancer, rare inherited disorders, asthma, neurological diseases, obesity, sexual dysfunction, hypertension, and the like). In some cases, the investigator will have identified the mutation and/or correlation in a very small population sample (e.g., in a single individual). The investigator may wish to determine the allele frequency of the SNP in the general population and may wish to generate an accurate diagnostic test to determine if an individual possesses the SNP, and is therefore at a higher risk than the general population of contracting or exhibiting the correlated disease or condition. In other embodiments, an investigator may have a SNP that is only suspected to correlate to a

disease state, and may wish to generate an accurate diagnostic test to screen large numbers of individuals who have been assessed for the presence or absence of the disease state in order to determine the whether the suspected correlation in fact exists. In other cases, the investigator may wish to determine the frequency of an allele within one or more populations for purposes including assessing risk for correlated disease states in the one or more populations. To address these needs, the investigator employs the systems and methods of the present invention.

The investigator uses a computer system to access a computer system of the service provider. In some embodiments, the investigator simply uses a personal computer system to access a publicly available Web site of the service provider. As discussed in Section I, above, the user transmits the identified target sequence containing the SNP to the computer system of the service provider. The target sequence is then processed through the in silico analysis systems and methods (Section I) and the detection assay design systems and methods (Section II) of the present invention. A report is sent to the investigator indicating any problems identified in the in silico analysis or design process and, in some embodiments, alternate target sequence suggestions are provided. The report may also indicate several options for the design of a detection assay from which the investigator may select. In some embodiments, at the time the original target sequence is submitted by the investigator, the investigator selects options for determining whether a report is provided (e.g., as opposed to simply proceeding with production without generating a report), the conditions under which a report is provided, and the information content of the report.

Once a target sequence is selected and design parameters for the detection assay components are selected (e.g., type of target [RNA or DNA] sequences of probes and primers, reaction temperatures, buffer conditions, etc.), information is passed to the production component of the systems and methods of the present invention (Section III). Production of the detection assay is carried out and quality control steps are used to ensure that the detection assay functions as intended (i.e., is capable of detecting the SNP in a sample). In some embodiments, the produced detection assay is screened against a plurality of known sequences designed to represent one or more population groups, *e.g.*, to determine the ability of the detection assay to detect the intended target amongst the diverse alleles found in the general population. Produced

assays are then shipped to detection assay users (e.g., the investigator who entered the target sequence and other investigators).

At each of the stages described above, information is tracked and stored. For example, the original target sequence request from the investigator is assigned a tracking number and information about the investigator (e.g., previous request information), information obtained from in silico analysis, information obtained from design analysis, and information obtained from production analysis (e.g., allele frequency information) is collected, correlated to the tracking number, and incorporated into the databases of the present invention. For example, allele frequency information is stored in a SNP allele frequency database, information obtained from in silico analysis and design analysis are stored for use in improved analysis of future target sequences, and information about investigators requesting the produced detection assays are stored and used to generate an information template for receiving detection assay data from the user after the assays are used (Section V). If in silico analysis determines that a SNP was previously characterized, the new request is assessed to see if it provides any additional information (*e.g.*, additional information provided by the new user), and such new information is integrated into the existing records for that SNP in the databases (*e.g.*, association databases, allele frequency databases). In some embodiments, the information about the target sequence and SNP obtained from the in silico, design, and production analysis are integrated with the information template to allow the investigator to access information (e.g., disease associations, allele frequency, etc.) prior to, during, or following use of the detection assay (e.g., information may be linked to a plate viewer function described in Section IV above).

The investigator uses the detection assay on one or more samples, *e.g.*, as described in Section V, above. Information and data are collected and returned to the systems of the service provider. Information and data obtained by the service provider from use of the detection assay are used for obtaining regulatory approval of clinical products corresponding to successful detection assays and to supplement information databases and improve in silico analysis, assay design, assay production, and future information dissemination to investigators. For example, additional allele frequency information may be obtained from the investigator. This information is used to supplement allele frequency databases. This information may also be used to increase or decrease the number of samples used during production analysis of allele frequency, as certain

samples (e.g., samples from particular ethnic groups, disease states, etc.) may be determined to be of limited information content (e.g., redundant) while others represent important, but previously unidentified or unappreciated populations for future analysis of allele frequency testing. Failure data from investigators (e.g., the failure of hybridization probes to hybridize to target sequences in a sample) is used in future in silico and design analysis.

As is clear from the above description, wide-scale use of the systems and methods of the present invention provides solutions to the unmet needs of the fields of bioinformatics and molecular diagnostics and medicine. Each phase of the invention, from target sequence validation and assay design and production to assay use and data collection provides a continuous circle of data generation and improvement. Wide scale use of the systems and methods of the present invention provides for the generation of reliable detection assays for the detection of any target sequence, wherein assays are designed to work for all individuals (e.g., a single assay that works for all individuals or a plurality of assays, each working for a known sub-set of the population). Databases generated using the systems and methods of the present invention provide comprehensive information pertaining to the allele frequency of mutations in one or more populations and the correlations of sequences and gene expression patterns to phenotypes. Thus, in some embodiments, the present invention provides detection assays and corresponding information databases and analysis systems for accurately screening entire populations (e.g., screening all human newborns) for sequences and expression patterns corresponding phenotypes (e.g., disease states, drug responses, etc.). Using the databases of the present invention, a specific sequence, combination of sequences, or expression patterns in an individual may be correlated to proven responses appropriate for the individual (e.g., avoidance of allergens, therapeutic drug treatments, gene therapy, preventive routes or behaviors, etc.).

B. Development of Clinical Detection Assays

As discussed above, of the thousands of markers evaluated using the systems and methods of the present invention, a sub-set of the markers are reliably detected by the detection assays of the present invention. Where a detection assay is shown to reliably detect a marker (e.g., a medically-relevant marker), detections assays for use as analyte-specific reagents or clinical diagnostics are prepared. Analyte-specific reagents and clinical diagnostics are regulated

in the United States. Using the systems and methods of the present invention, data generated during the development of the detection assays is used to support regulatory approval of the detection assay for used as analyte-specific reagents and clinical diagnostics. Because the present invention provides easy-to-use, efficient, accurate detection assays (e.g., the INVADER assay) that can be produced for thousands of unique markers at high production capacity and because the present invention provides systems and methods for widespread testing and data collection of thousands of samples with each of the thousands of unique detection assays, sufficient information is gathered to support regulatory approval of numerous clinical products. The present invention provides systems and methods for testing all identified markers, selecting markers that are suitable for clinical use, and collecting data in support of regulatory approval for every clinically relevant marker. The specific regulatory requirements for analyte-specific reagents and in vitro diagnostics are outlined below.

A major class of markers and mutations that find use in diagnostics are drug metabolism enzymes. Drug-metabolizing enzymes (DMEs) help the body to break down drugs properly and enable their therapeutic effects. One or more variations in a DME gene may affect how a person responds to a particular drug. As a result, one person may respond positively to a drug, while another may suffer adverse reactions to the same drug and still another will be unaffected by it. Detection assays that detect DME mutations expand the markets of existing drugs and the revival of drugs not allowed to or removed from the market because of adverse drug reactions or lack of therapeutic effect. The use of the present invention also provides high throughput screening of prospective new drug compounds that can eliminate potentially toxic drug candidates from development early in the process; reduces the cost and risk of clinical drug trials through pre-trial genetic screening; and provides clinical diagnostics to determine appropriate drug and dosage before prescription to avoid adverse drug reactions.

I. Adverse Drug Reactions and Genetic Variation

More than 3 billion prescriptions are written each year in the U.S. alone, effectively preventing or treating illness in hundreds of millions of people. But prescription medications also can cause powerful toxic effects in a patient. These effects are called adverse drug reactions (ADR). Adverse drug reactions can cause serious injury and or even death. Differences in the

ways in which individuals utilize and eliminate drugs from their bodies are one of the most important causes of ADRs (MedWatch).

More than 106,000 Americans die – three times as many as are killed in automobile accidents – and an additional 2.1 million are seriously injured every year due to adverse drug reactions. ADRs are the fourth leading cause of death for Americans. Only heart disease, cancer and stroke cause more deaths each year. Seven percent of all hospital patients are affected by serious or fatal ADRs. More than two-thirds of all ADRs occur outside hospitals. Adverse drug reactions are a severe, common and growing cause of death, disability and resource consumption in North America and Europe.

ADRs most commonly occur when the body cannot change a drug quickly enough into a form that it can use and then eliminate. A drug compound goes through a series of many changes as it is being processed in the body, some of which actually may make the drug more toxic before it is changed again. If this toxic form of the drug is not changed or eliminated by the body, it can cause illness, permanent liver damage or even death. Proteins called drug-metabolizing enzymes (DMEs) make these changes as the body processes a drug.

All drugs have the potential to cause ADRs. The most common, however, are central nervous system agents (antidepressants, anticonvulsants, eye and ear preparations, internal analgesics and sedatives), anti-infectious drugs (penicillin and the sulfa antibiotics), anti-cancer drugs and cardiovascular drugs cause the most ADRs. Cardiovascular drugs alone cause 25 percent of all ADRs.

It is estimated that drug-related anomalies account for nearly 10 percent of all hospital admissions. Drug-related morbidity and mortality in the U.S. is estimated to cost from $76.6 to $136 billion annually.

A. Cytochrome p450 polymorphisms

The cytochrome p450 (CYP) superfamily comprises a group of enzymes that play an essential role in the bio-transformation of medically relevant compomounds. Approximately 40% of CYP isoforms are polymorphic, including CYP1A2, 3A4, 2B6, 2CP, and 2C19 (see also Table 8 below). Accurate genotyping of patients for these and other p450 loci is important because allelic variants may lead to loss of efficacy or toxic accumulation. These consequences

are particularly pronounced in the perioperative interval with multiple low therapeutic ratio substrates competing for shared CYP pathways.

TABLE 8

Gene	Location	Substrate
CYP1AI	15q22-q24	Benzo(a)pyrene, phenacetin
CYP1A2	5q22-qter	Acetaminophen, amonafide, caffeine, paraxanthine, ethoxyresorufin, propranolol, fluvoxamine
CYP1B1	2p21	estrogen metabolites
CYP2A6	19q13.2	Coumarin, nicotine, halothane
CYP2B6	19q13.2	Cyclophosphamide, aflatoxin, mephenytoin
CYP2C19	10q24.1-24.3	Mephenytoin, omeprazole, hexobarbital, mephobarbital, propranolol, proguanil, phenytoin
CYP2C8	10cen-q26.11	Retinoic acid, paclitaxel
CYP2C9	10q24	Tolbutamide, warfarin, phenytoin, nonsteroidal anti -inflammatories
CYP2D6	22q13.1	Flexainide, guanoxan, methoxyamphetamine, N -propylajmaline, perhexiline, phenacetin, phenformin, propafenone, sparteine
CYP2E1	10q24.3-qter	N-Nitrosodimethylamine, acetaminophen, ethanol
CYP3A4/3A5/3A7	7q21.1	Macrolides, cyclosprorin, tacrolimus, calcium channel blockers, midazolam, terfenadine, lidocaine, dapsone, quinidine, triazolam, etopside, teniposide, lovastatin, tamoxifen, steroids, benzo(a)pyrene

One example of a drug influenced by a CYP loci is the drug WARFARIN, which is a blood thinner routinely prescribed to prevent or treat blood clots, especially those associated with heart attack or heart value replacement and to reduce the risk of death, another heart attack or stroke after a heart attack. More than 19 million prescriptions for the drug were written in 2000. Approximately eight percent of whites and two percent of blacks have a genetic variation

(CYP2C9*3) that causes the body to slow its metabolism of WARFARIN, which can cause bleeding that can resulting in the loss of large amounts of blood.

Genetic screening for this variation allows health care professionals to prescribe the correct dosage of WARFARIN to avoid the severe bleeding and to preclude the use of aspirin, which could further thin the blood and amplify the adverse reaction.

Many of the p450 genes are highly polymorphic. INVADER assays can be used to detect particular polymorphisms in p450 genes in order to help prevent adverse drug reactions in patients. One example is the CYP2D6 gene. Figure 82 shows the various polymorphisms for this gene. Importantly, the two CYP2D pseudogenes, CYP2D7 and CYP2D8, share many of the identified polymorphisms of CYP2D, and over 80% sequence similarity. Therefore, to prevent false positive results, due to detection of the two psuedogenes, a CYP2D6 specific Triplex PCR amplification reaction was developed to integrate with the INVADER assay. The three PCR products are amplified from genomic template in a single tube using CYP2D6 specific PCR primers with a 35 cycle PCR reaction of 95 degrees Celsius for 20 seconds and 68 degrees Celsius for 2 minutes (see Figure 83).

Next, a 1/20 dilution of the CYP2D6 specific PCR products are used as a template for polymorphism detection using the Biplex INVADER assay system in a single well of a 96 or 384 well plate. Two serial INVADER assay reactions occur simultaneously, target detection and allele discrimination takes place in the primary INVADER reaction, while signal amplification takes place in the secondary INVADER reaction using a set of universal signal probes. The entire assay is isothermal and only requires a single step to set up. In addition to this, signal can be read and alleles called after only 20 minutes incubation at 63 degrees Celsius following an initial 5 minutes 95 degrees Celsius denaturation step (See Figure 84). The results of a screen of 175 individuals using this approach is shown in Figures 85 and 86.

B. Detection Assays and Drugs

Most prescription drugs are currently prescribed at standard doses in a "one size fits all" method. This "one size fits all" method, however, does not consider important genetic differences that give different individuals dramatically different abilities to metabolize and derive benefit from a particular drug. Genetic differences may be influenced by race or ethnicity

(See Figure 87). As such, certain groups of people considered at high risk (e.g. for an adverse drug reaction) are tested with a detection assay prior to administration of the drug. Also, detection assays (e.g. in panels) to identify which classes of patients will likely receive benefit from a candidate drug being developed.

If a health care provider knows both which genetic markers in particular DMEs are important for a particular drug and which variations of those genetic markers a patient has, it will be significantly easier to avoid dangerous ADRs. The genetic diagnostic panels of DME variations provided by the present invention allow one to determine the best course of treatment for each patient and to prescribe the most appropriate drug at the safest dosage, all based on an simple, easy-to-use assessment of the patient's unique genetic make-up.

Genetic markers for drug-metabolizing enzymes (DMEs) have enormous potential for dramatically altering the process that determines not only whether a drug enters the market, but also whether a drug that has been withdrawn can be "revitalized." Individual responses to a particular drug often arise from variations within the genes that produce DMEs. An understanding of which DMEs are involved with helping the body eliminate a particular drug will be coupled with the knowledge of variations cause the body to metabolize the drug too quickly or too slowly. This important medical insights forms the foundation for high-resolution genetic diagnostic panels of thousands of DME variations that find use by health care providers before prescribing a particular drug. Those found to have genetic variation(s) associated with an adverse response to a particular drug are prescribed a different drug, one that is safe for them. Patient safety is enhanced significantly and those in desperate need of the therapeutic effects of a drug that has been withdrawn from the marketplace once again have access to an effective medication.

The development of a single new drug is estimated to cost $500 million, with much of the expense being incurred in the final phases. The use of DME markers of the present invention increases the efficiency of drug development in every phase, but is particularly useful in eliminating potentially toxic compounds from development in the earliest phases, before the majority of development dollars have been spent. Even after the expense of development, it is estimated that the most commonly used drugs will be effective in only 30-60 percent of patients with the same illness or disease. DME markers are used during drug development for the

parallel development of genetic diagnostics that are administered at the point of care to avoid adverse drug reactions and improve the effectiveness of the drug. Thus, the present invention improves target discovery (the identification of new drug targets), preclinical toxicity determinations (the elimination of compounds that might cause ADRs early in the development process), lead compound prioritization (the prioritization of potential new drug compounds that have the desired effect and show no potential for ADRs), and clinical trial patient stratification (the ability to select potential participants with similar DMEs for clinical studies).

Representative drugs that have been withdrawn from the market since 1997 are shown in Table 9.

TABLE 9

Withdrawn	Clinical Name	Reason for Using	ADR
2001	Cerivastatin	Cholesterol control	Muscle cells damage
2001	Repacuronium bromide	Muscle relaxant	Breathing problems
2000	Alosetron hydrochloride	Spastic colon	Liver damage
2000	Cisapride	Heartburn	Heartbeat problems
2000	Troglitizone	Type 2 diabetes	Liver damage
1999	Astemizole	Allergies	Heart problems
1998	Bromfenac	Pain relief	Liver damage
1998	Mibefradil	High blood pressure	Drug interactions
1997	Fenfluramine and	Obesity	Heart valve damage
1997	Phentermine	Obesity	Heart valve damage

II. Analyte-Specific Reagents

In some embodiments, components of nucleic acid detection assays are sold as analyte specific reagents (ASRs). ASRs are restricted devices under section 520(e) of the Federal Food, Drugs, and Cosmetic Act and 21 CFR 809.30 and are subject to specific restrictions. ASRs may only be sold to "in vitro diagnostic manufacturers": clinical laboratories regulated under the

Clinical Laboratory Improvement Amendments of 1988 (CLIA), as qualified to perform high complexity testing under 42 CFR part 493 or clinical laboratories regulated under VHA Directive 1106 (available from Department of Veterans Affairs, Veterans Health Administration, Washington, DC 20420); and organizations that use the reagents to make tests for purposes other than providing diagnostic information to patients and practitioners (e.g., forensic, academic, research, and other nonclinical laboratories). In addition, ASRs must be labeled in accordance with Sec. 809.10(e). Advertising and promotional materials for ASRs must include the identity and purity (including source and method of acquisition) of the analyte specific reagent and the identity of the analyte; the statement for class I exempt ASR's: "Analyte Specific Reagent. Analytical and performance characteristics are not established"; include the statement for class II or III ASR's: ``Analyte Specific Reagent. Except as a component of the approved/cleared test (name of approved/cleared test), analytical and performance characteristics are not established"; and must not make any statement regarding analytical or clinical performance.

Any laboratory that develops an in-house test using the ASR is required to inform the ordering person of the test result by appending to the test report the statement: ``This test was developed and its performance characteristics determined by (Laboratory Name). It has not been cleared or approved by the U.S. Food and Drug Administration." This statement would not be applicable or required when test results are generated using the test that was cleared or approved in conjunction with review of the class II or III ASR. Ordering in-house tests that are developed using analyte specific reagents is limited under section 520(e) of the act to physicians and other persons authorized by applicable State law to order such tests.

III. In vitro Diagnostic Detection Assays

In some embodiments, assays for detecting genetic variation are marketed as in vitro diagnostic tests. The marketing of such kits in the United States requires approval by the Food and Drug Administration (FDA). The FDA classifies in vitro diagnostic kits as medical devices. As such, the pre-market applications for most in vitro diagnostics are submitted to the FDA under the 510(k) regulations and are referred to as 510(k) applications. The 510(k) regulations specify categories for which information should be included.

Each person who wants to market Class I, II and some III devices intended for human use in the U.S. must submit a 510(k) to FDA at least 90 days before marketing unless the device is exempt from 510(k) requirements. Classification of devices are determined by finding the regulation number that is the classification regulation for each device. This can be accomplished searching the classification database for a part of the device name, or, if the device panel (medical specialty) to which the device belongs is known, going directly to the listing for that panel and identify the device and the corresponding regulation. Links to both database can be found on the web page of the FDA.

A 510(k) is a premarketing submission made to FDA to demonstrate that the device to be marketed is as safe and effective, that is, substantially equivalent (SE), to a legally marketed device that is not subject to premarket approval (PMA). Applicants must compare their 510(k) device to one or more similar devices currently on the U.S. market and make and support their substantial equivalency claims. A legally marketed device is a device that was legally marketed prior to May 28, 1976 (preamendments device), or a device which has been reclassified from Class III to Class II or I, a device which has been found to be substantially equivalent to such a device through the 510(k) process, or one established through Evaluation of Automatic Class III Definition. The legally marketed device(s) to which equivalence is drawn is known as the "predicate" device(s).

Applicants must submit descriptive data and, when necessary, performance data to establish that their device is SE to a predicate device. The data in a 510(k) is to show comparability, that is, substantial equivalency (SE) of a new device to a predicate device. A claim of substantial equivalence does not mean the new and predicate devices must be identical. Substantial equivalence is established with respect to intended use, design, energy used or delivered, materials, performance, safety, effectiveness, labeling, biocompatibility, standards, and other applicable characteristics.

Once the device is determined to be SE, it can then be marketed in the U.S. If the FDA determines that a device is not SE, the applicant may resubmit another 510(k) with new data, file a reclassification petition, or submit a premarket approval application (PMA). The SE determination is usually made within 90 days and is made based on the information submitted by the applicant.

A 510(k) is required when introducing a device into commercial distribution (marketing) for the first time, when proposing a different intended use for a device which is already in commercial distribution, and when there is a change or modification of a device already marketed that could significantly affect its safety or effectiveness.

Information required in an application under 510(k) includes:

1) The in vitro diagnostic product name, including the trade or proprietary name, the common or usual name, and the classification name of the device.
2) The intended use of the product.
3) The establishment registration number, if applicable, of the owner or operator submitting the 510(k) submission; the class in which the in vitro diagnostic product was placed under section 513 of the FD&C Act, if known, its appropriate panel, or, if the owner or operator determines that the device has not been classified under such section, a statement of that determination and the basis for the determination that the in vitro diagnostic product is not so classified.
4) Proposed labels, labeling and advertisements sufficient to describe the in vitro diagnostic product, its intended use, and directions for use. Where applicable, photographs or engineering drawings should be supplied.
5) A statement indicating that the device is similar to and/or different from other in vitro diagnostic products of comparable type in commercial distribution in the U.S., accompanied by data to support the statement.
6) A 510(k) summary of the safety and effectiveness data upon which the substantial equivalence determination is based; or a statement that the 510(k) safety and effectiveness information supporting the FDA finding of substantial equivalence will be made available to any person within 30 days of a written request.
7) A statement that the submitter believes, to the best of their knowledge, that all data and information submitted in the premarket notification are truthful and accurate and that no material fact has been omitted.
8) Any additional information regarding the in vitro diagnostic product requested that is necessary for the FDA to make a substantial equivalency determination. A request for additional information will advise the 510(k) submitter that there is insufficient

information contained in the original 510(k) submission for a substantial equivalent determination to be made. In this situation the 510(k) submitter may: (a) submit the requested data or a new 510(k) containing the requested information, or (b) submit a PMA application in accordance with section 515 of the FD&C Act. If the additional information is not submitted within 30 days following the date of the request, the FDA may consider the 510(k) to be withdrawn.

Factors used by FDA reviewers in determining substantial equivalency include:

1) Does the in vitro diagnostic device have the same intended use as a currently marketed device (sometimes referred to as a "predicate device"), e.g., nucleic acid diagnostic assay?
2) Does the in vitro diagnostic device have the same technological characteristics, e.g., nucleic acid probes?
3) If new technological features are present, e.g., DNA probe, monoclonal antibody, do they raise new questions regarding safety and effectiveness?

Additionally, the following questions will be used by FDA reviewers to assess whether an in vitro diagnostic device that includes technological changes is substantially equivalent to a predicate device.

1) Does the in vitro diagnostic device pose the same type of questions about safety and effectiveness as the predicate device?
2) Are there accepted scientific methods for assessing the impact of technological changes on safety and effectiveness, e.g., accuracy, specificity, sensitivity, precision?

Data generated using the system and methods of the present invention provides sufficient information to obtain approval on the detection assays. Prior to the present invention, only a small number of in vitro diagnostic detection assays have been approved. The present invention provides system and methods for producing approved detection assays for the hundreds of most medically relevant markers. As such, the present invention provides the predicate devices for many markers by which future detection assays will be compared. In some embodiments, the

present invention provides methods for obtaining regulatory approval of new detection assays by comparing data obtained with the new detection assay (e.g., data obtained using the systems and methods of the present invention) to a predicate device obtained by using the systems and methods of the present invention.

IV. Product Development

The present invention provides systems, computer programs, graphical user interfaces, and methods for ordering, manufacturing, and delivering detection assays. In some preferred embodiments, an electronic detection assay ordering system is provided to facilitate the utilization of systems and methods for acquiring and analyzing biological information (e.g., systems and methods for developing detection assays and for use of detection assays in basic research discovery to facilitate selection and development of clinical detection assays).

The discovery of a new gene sequence suspected of correlating to a disease condition offers a starting point for understanding the correlation and hopefully, of leading to a treatment for the condition. This data is input into the one or more components of the system of the present invention. However, extensive amounts of work need to be conducted before a useful and safe treatment can be obtained. The systems and methods of the present invention provide an efficient and thorough means to accelerate the time between initial discovery and useful treatment, and provide the tools for diagnosis and development of therapies using components of a production facility that provides for the efficient ordering, production, and shipment of detection assays. Prior to the invention there was no way for a researcher or other user to determine if a detection assay was commercially available for a SNP of interest so that research could be conducted. For example, where a mutation (e.g., a single nucleotide polymorphism; "SNP") is suggested to correlate with a disease, the present invention provides systems for identifying an optimal target sequence from which an assay is developed to detect the presence of the mutation in a sample. The present invention also provides systems and methods for designing and producing a highly accurate detection assay or other detection assays directed to the optimized target sequence. The assay may then be used to detect the mutation in a large number of samples to determine the accuracy of the original proposed correlation and to determine additional information about the mutation (e.g., the allele frequency of the mutation in

any desired population, data necessary for obtaining approval for clinical products from regulatory agencies, etc.). Data collected from these experiments is then analyzed and processed by systems and methods of the present invention to facilitate improved target selection, the identification of additional mutations, the identification of additional correlations, and the design of clinical assays for diagnosing the presence of the mutations in subjects (*e.g.*, to identify subjects that are appropriate candidates for a particular type of therapy). All of this data is fed to various components of the invention.

In some embodiments of the present invention, efficient, sensitive detection assays are provided. The assays are used by users (e.g. researchers) to collect test result data from a plurality of samples. Data obtained from the samples is used, among other purposes, to validate the detection assay (e.g. data is returned to the databases of the data management systems of the present invention). Validated data is then fed to the various components of the invention. For example, collected test result data is used to provide evidence necessary to support approval (e.g., FDA approval) of clinical products corresponding to the detection assay, and can be fed to and stored on a database which is a part of one or more components of the invention. In some embodiments, a plurality of detection assays are combined into a panel and the panels are used to simultaneously collect data for multiple genetic markers. The collected data is used to provide evidence necessary to support approval of clinical products corresponding to one or more of the detection assays on the panel, and can be sent from a remote site or sites to any of the components of the present invention for optimization of a detection assay or production thereof. In some embodiments, a party provides detection assays at a reduced cost, at a subsidized cost, or at no cost to users (e.g. researchers), and data collected by the users is used to support development and/or approval of clinical detection assay products by the providing party and is fed to a database that is linked to one of the components of the present invention. In other words, detection assays are produced (e.g. by the methods described above), and shipped to a user a reduced charge in exchange for detection assay result data (e.g. returned to one or more databases of the data management systems of the present invention via the internet). The result data is then used to forecast demand for a certain assay, reagent production need. In yet another

variant, the data is fed to the inventory component so that inventory of a particular assay or panel can be regulated, (e.g. increased or decreased accordingly).

In some embodiments, the present invention provides systems, routines and methods for the development of research and clinical diagnostic products using a multi-step process (*i.e.* product development funnel) and data related thereto. A schematic summary of such a process is shown in Figure 88. This figure shows four stages of detection assay development from discovery-based detection assays (e.g., identification and characterization of sequences and mutations), to medically associated marker detection assays (e.g., detection assays directed to markers associated directly or indirectly with one or more medically important conditions), to analyte-specific reagent assays, to clinical diagnostic detection assays (e.g., in vitro detection of established clinical markers). The funnel shown in Figure 88 represents the fact that a large number of markers may be examined in the discovery phase, leading to a sub-set that are appropriate for each of the subsequent phases It is appreciated that detection assay development utilizes databases that form a part of one or more components hereof. A discovery-based detection assay data or designation is correlated to a first group of detection assays and stored on a database, and utilized with routines of various components of the invention. Medically associated marker data or designations for another group of detection assays are stored and utilized in routines associated with components of the invention. ,The same holds true for analyte-specific reagent data or designations for detection assays and clinical diagnostic data or designations for various detection assays. This data is used in the manufacturing, pricing and inventory processes and routines described herein.

The following section describes how DNA analysis products directed to SNP detection are moved through the funnel. The focus on DNA products and SNP detection is for clarity only. RNA analysis products and other analysis products also find use in the present invention (e.g., for detecting and quantitating gene expression and other RNA levels using the same product strategy, including detection of splice variants and polymorphism variants). Figure 89 shows a schematic summary of the discovery phase. In this phase, detection assays or one or more variety are directed to the thousands to hundreds of thousands of markers are generated. This data is stored on databases of various components thereof for use in the production processes and web order entry routines and processes described herein. While the association of

certain SNPs to particular medical conditions has been determined, association has not been established for the majority of SNPs. The present invention provides a broad menu of assays and assay data that is presented to a prospective customer for purchase. For example, more than 80,000 unique assays applying the INVADER assay technology (Third Wave Technologies, Madison, WI) have been developed, manufactured and shipped for genotyping research to associate specific SNPs with predisposition to disease. Many of the assays have been sent to collaborative customers at low cost in exchange for access to collected data and rights to commercialize discoveries made with these collaborators.

Figure 90 shows a schematic summary of the "Medically Associated" phase. Detection assay data is correlated to medically associated data and stored on storage device communicatively linked to one or more components of the invention. As use of detection assays reveals the potential association of a SNP with a medical condition, it is designated a potential clinical marker and earmarked for inclusion on one or more Medically Associated Panels (e.g., panels comprising a plurality of detection assays directed at two or more distinct markers). This data is used in one or more components of the invention for production or pricing. Using this approach, the association of certain SNPs has been established and panels have been prepared. Detection assays are added for new makers to panels as those markers are associated and moved down the funnel. Figure 90 shows two types of panels created using the systems and methods described herein, those containing markers specific to certain disease types or fields (e.g., cardiovascular disease, oncology, immunology, metabolic disorders, neurological disorders, musculoskeletal disorders, endocrinology, and other genetic diseases) and large panels (e.g., containing 10 thousand or more markers) directed to all known medically relevant diseases. It is appreciated that data of detection assays for these various disease types are correlated, stored on databases, and used in the production processes and web user interface described herein.

In one variant, researchers using the panels validate the associations of particular genetic markers to specific medical conditions Analyte-Specific Reagents (ASRs) phase). Once an association is valid, the assay is moved one step further down the funnel and, more importantly, into the clinical market. At this point a price point may change for the assay, and appropriate price data points are correlated to other detection assay data. The ASR format permits the use of the assay in clinical settings without full FDA approval as the user, a certified clinical laboratory,

validates the assay for the particular use. The format also allows for the generation of demand and the monitoring of demand using routines and data for a clinical marker or set of markers prior to deciding to seek FDA approval to market it as a in-vitro diagnostic tool (See Figure 91).

In yet another variant, which may include a Diagnostics phase, once sufficient market demand exists for a particular assay, full regulatory approval is sought to market the assay as an *in vitro* diagnostic (IVD). While IVD products are represented as occupying the smallest part of the funnel, they are the largest potential revenue source, as shown schematically in Figure 92. At this point new or higher price point data may be correlated to one or more components of the detection assay data. As a detection assay is moved from research to clinical use, the cost to produce it does not increase significantly, while the revenue and profit margin it generates increase exponentially. The assay manufactured and shipped as an IVD is fundamentally the same assay that entered the top of the funnel as a discovery tool (although improvements or changes may be made during the process, as described below).

Examples of products for each of the funnel phases is shown in Figure 91 for both genotyping and SNP detection of DNA samples (e.g., samples containing genomic DNA) and expression analysis. For the discovery phase, the systems and methods of the present invention have been applied to generate over 80 thousand unique SNP detection assays with the ability to add six to ten thousand, or more, additional unique SNP detection assays per month. In some embodiments, discovery panels are manufactured using the methods and systems herein that are directed to SNP analysis of entire genes or chromosomes. The present invention also provides systems and methods for custom design of detection assays at any phase of the funnel (i.e., custom design of research and clinical detection assays) by an end user or internally at a production facility. For the medically associated phase, specific panels have been developed for DNA analysis and a large number of expression analysis detection assays have been developed or are in development. For custom panels, customers may elect one or more markers of their choosing for use on the panel and input this data from a catalogue or markers presented on the customer order componend. In some embodiments, customers enter their desired panel components into a user interface of a software program and the received data is sent for analysis and production to one or more components of the invention.

In some embodiments, the funnel process is facilitated by a low cost, easy-to-use assay (e.g., the INVADER assay) and a production process that allows substantial numbers of detections assays to be generated using the methods, routines and systems of the present invention. Such assays provide the necessary features (e.g., accuracy, sensitivity, ease-of-use, amenability to high throughput automated analysis, etc.) to allow wide-spread use by researchers, such that sufficient data is collected to process large numbers of detections assays through the funnel process. Widespread data collection results in the assay becoming a standard for use in discovery of the genetic basis of disease and management of personalized medicine strategies. For example, the present invention provides systems and methods to allow regulatory approval of clinical diagnostic products of every suitable marker. Detection assays for which regulatory approval is sought have detection assay data correlated with a regulatory approval designation or data, and may be processed using the systems and methods described herein in a manner that is different from, for example, RUO assays. These assay may undergo more rigorous quality control processes described herein.

In certain embodiments, a disease associated assay for a particular type of condition (e.g. Cardiovascular, DMD, CF, oncology, etc.) is sought to be developed. Disease condition data by be correlated with SNP data or RNA data or detection assay data. This correlated data is then used in one or more components of the present invention. Figure 93 shows an approach that may be used to develop particular disease associated assays. The approach shown in Figure 93, or similar approaches, shows how a pool of medically associated SNP assays is first identified (e.g. by the systems of the present invention that allow results of assay use to be collected and analyzed), and then this pool is further processed to develop commercial products. In particular, figure 93 shows a Medically Associated Panel (MAP) development track and a Clinical development track, how particular assays move throught the development process, how failed assays are further developed, and how successful assays are marketed (e.g. first as Reasearch Use Only (RUO) assays, and then launed as ASRs and/or in vitro diagnostics (IVD)).

In some embodiments, the present invention provides an ASR fast track development process. One of the barriers to a rapid and facile ASR product development lies in the relatively lengthy time required for some of the candidate ASR's to be researched and developed. The period from identification of an ASR to the time that validation studies can begin has ranged

from several months to years. However, the integrated systems and processes of the present invention allow this process to be sped up dramatically.

The rapid identification and evaluation of candidate ASRs may, for example, occur in several stages. Overview of the ASR fast track is presented in Figure 71. The first step in the process is the identification of "Super SNPs". Super SNPs are generally those SNPs and/or detection assays that have extraordinary performance characteristics from an aggregate of SNPs or detection assays that have been designed and tested. In preferred embodiments, a screening process like the one shown in figure 95 is employed. Preferably, a production databases (including QC performance data) of previously designed and tested SNP assays is employed as the starting point. Using a production database as the starting point has many advantages. For example, the SNPs within the database already are likely to have some importance as they have been chosen by a customer (optionally at the customer order entry component of the invention). Also, employing the QC performance data within the database as an initial screen generally eliminates the need for further development.

Once a Super SNP or set of Super SNPs has been identified, the relevance of the SNP site as an Analyte Specific Reagent (ASR) is then determined. This may be done using databases (e.g. public databases, and those on an internal data management system, see above) and routines to compare the target region of the Super SNPs to these databases. If this database search indicates that this target region has relevance to any number of markets (e.g. clinical ASR and/or reasarch use only ASRs) that SNPs status is changed from Super SNP to ASR/RUO candidate on a database used herein.

Next a market review is performed (see Figure 94). For example, using market research information, ASR/RUO Candidate products are further evaluated as to which market this candidate is most appropriate. Appropriate designations are made correlating this data to detection assay data. Once an ASR/RUO Candidate has been evaluated as to the proper market area, validation studies are performed.

V. Panels, Libraries, Databases

The present invention provides methods and compositions for treating nucleic acid, and in particular, methods and compositions for detection and characterization of nucleic acid

sequences and sequence changes. In particular, the present invention provides detection assay panels comprising an array (e.g. microarray) of different detection assays. The arrays are manufactured using the systems and methods described herein. The detection assays include assays for detecting mutations in nucleic acid molecules and for detecting gene expression levels. Assays find use, for example, in the identification of the genetic basis of phenotypes, including medically relevant phenotypes and in the development of diagnostic products, including clinical diagnostic products. The present invention also provides systems and methods for data storage, including data libraries and computer storage media comprising detection assay data.

As discussed above, the present invention is not limited by the nature of the detection assays used in the panels or microarrays of the present invention. A wide variety of available detection technologies find use with the present invention, including those described in detail herein. Purely for illustration purposes, much of the disclosure herein, highlights the use of panels with the INVADER assay detection system (Third Wave Technologies, Madison, WI). In particular, the following description provides a detailed analysis of how to apply a detection assay technology (e.g., the INVADER assay) to the systems and methods of the present invention. One skilled in the art will appreciate the applicability of the invention to other detection technologies.

The panels and microarrays of the present invention mark a significant advancement in genetic variation analysis products, allowing researchers to genotype many (e.g., hundreds to thousands) of genetic variations simultaneously in a simple, easy to use, ``just add DNA" format. For example, the present invention provides panels comprising a plurality of different INVADER assay detections assays on a single panel. Such panels comprise, for example, the detection assay described in Figure 96 which include tests for single nucleotide polymorphisms (SNPs) and other mutations that have been associated with diabetes, asthma, deafness, hypertension and other medically relevant conditions.

The panels of the present invention enhance the medical community's ability to detect, catalog and utilize clinically relevant mutations. The availability of disease specific, ready to use panels not only facilitate the additional clinical research needed to extend the initial findings of medical association, but also establish the clinical utility of specific genetic variation analysis products, helping to accelerate their ultimate use and sale as diagnostic tools to the clinical

market. Data of which detection assays are part of a respective panel are stored on databases that optionally form part of the components herein and are utilized in the various components of the invention for product presentation, production, inventory control, billing and shipping.

In some embodiments, panels comprise detection assays that allow for simultaneous detection of multiple variations in a sample using identical reaction conditions. For example, the INVADER assay detection panels of the present invention enable scientists to detect multiple genetic variations in one individual using the same array (e.g., microtiter plate) because each well of the plate contains a different SNP or mutation test, all run under identical conditions.

In some preferred embodiments, panels are designed for ease of use. For example, the INVADER assay panels of the present invention are readily produced as products that can be shipped ready to use with stable, dried-down reagents in each reaction site on an array (e.g., each well in a microtiter plate). All the user must do is add genomic or amplified DNA to detect variations in a wide range of genes.

In some preferred embodiments, each detection assay on a panel allows for biplex or multiplex analysis. For example, the INVADER assay may be applied in a biplex format, which enables the simultaneous detection of all variations for each SNP. For example, the presence of the three possible genotypes for an A-C polymorphism -- AA, AC or CC -- can be determined in a single well. Since each well yields at least one positive signal -- A or C or both -- the biplex format also provides an internal control.

The panels of the present invention may also be used in conjunction with bioinformatics tools. For example, genetic variation analysis kits comprising the panels of the present invention and software that can be run on virtually all hardware platforms. The bioinformatics software couples the performance and ease of use of the panel product with a data collection and analysis tool. It transforms instrument readings into useful genetic variation data and links it to searchable background information about each detection assay SNP or mutation and additional information available through publicly available databases, including Johns Hopkins' Online Mendelian Inheritance in Man (OMIM) and NCBI's GenBank.

In some embodiments, information pertaining to the panels (e.g., design features, bioinformatics information, test result data, etc.) is collected and stored in one or more databases. Thus, the present invention provide detection assay libraries and searchable databases for use in

compiling and analyzing information and for selecting assays for use in future panels and for development of clinical detection assays.

In some embodiments, the panels of the present invention are in microarray format (e.g. oligonucleotdies are Data of which detection assays are part of a respective panel are stored on databases that optionally form part of the components herein and are utilized in the various components of the invention for product presentation, production, inventory control , billing and shipping. attached to a solid surface such that a detection assay may be perfored on the solid surface). In other embodiments, the solid support serves as a platform on which microwells are printed/created and the necessary reagents are introduced to these microwells and the subsequent reaction(s) take place entirely in solution. Creation of a microwells on a solid support may be accomplished in a number of ways, including; surface tension, and etching of hydrophilic pockets (e.g. as described in patent publications assigned to Protogene Corp.). For example, the surface of a support may be coated with a hydrophobic layer, and a chemical component, that etches the hydrophobic layer, is then printed on to the support in small volumes. The printing results in an array of hydrophilic microwells. An array of printed hydrophobic towers may be employed to create micorarrays. A surface of of a slide may be coated with a hydrophobic layer, and then a solution is printed on the support that creates a hydrophilic layer on top of the hydrophobic surface. The printing results in an array of hydrophilic towers. Mechanical microwells may be created using physical barriers, +/- chemical barriers. For example, microgrids such as gold grids may be immobolized on a support, or microwells may be drilled into the support (e.g. as demonstrated by BML). Also, a microarray may be printed on the support using hydrophilic ink such as TEFLON. Such arrays are commercially available through Precision Lab Products, LLC, Middleton, WI. In yet another variant, data of customer preferences with respect to the format of the detection assay array are stored on a database used with components of the invention. This information can be used to automatically configure products for a particular customer based upon minimal identification information for a customer, e.g. name, account number or password.

Many types of methods may be used for printing of desired reagents into microwell arrays. In some embodiments, a pin tool is used to load the array mechanically (see, e.g., Shalon, Genome Methods, 6:639 [1996], herein incorporated by reference). In other embodiments, ink

jet technology is used to print oligonucleotides onto a solid surface (*e.g.*, O'Donnelly-Maloney *et al.*, Genetic Analysis:Biomolecular Engineering, 13:151 [1996], herein incorporated by reference).

Examples of desired reagents for printing into/onto microwell arrays include, but are not limited to, molecular reagents, such as INVADER reaction reagents, designed to perform a nucleic acid detection assay (e.g., an array of SNP detection assays could be printed in the wells); and target nucleic acid, such as human genomic DNA (hgDNA), resulting in an array of different samples. Also, desired reagents may be simultaneously supplied with the etching/coating reagent or printed into/onto the microwells/towers subsequent to the etching process. For arrays created with mechanical barriers the desired reagents are, for example, printed into the resulting wells. In some embodiments, the desired reagents may need to be printed in a solution that sufficiently coats the microwell and creates a hydrophilic, reaction friendly, environment such as a high protein solution (e.g. BSA, non-fat dry milk). In certain embodiments, the desired reagents may also need to be printed in a solution that creates a "coating" over the reagents that immobilizes the reagents, this could be accomplished with the addition of a high molecular weight carbohydrate such as FICOLL or dextran.

Application of the target solution to the microarray (or reaction reagents if the target has been printed down) may be accomplished in a number of ways. For example, the solid support may be dipped into a solution containing the target. or putting the support in a chamber with at least two openings then feeding the target solution into one of the openings and then pulling the solution across the surface with a vacuum or allowing it to flow across the surface via capillary action. Examples of devices useful for performing such methods include, but are not limited to, Tecan – GenePaint system, and AutoGenomics AutoGene System. In yet another embodiment spotters commercially avialable from Virtek Corp. as used to spot various detection assays onto plates, slides and the like.

In some embodiments, solutions (e.g. reaction reagents or target solutions) are dragged, rolled, or squeegeed accross the surface of the support. One type of device useful for this type of application is a framed holder that holds the support. At one end of the holder is a roller/squeegee or something similar that would have a channel for loading of the target solution in front of it. The process of moving the roller/squeegee across the surface applies the target

solution to the microwells. At the end opposite end of the holder is a reservoir that would capture the unused target solution (thus allowing for reuse on another array if desired). Behind the roller/squeegee is an evaporation barrier (e.g., mineral oil, optically clear adhesive tape etc.) and it is applied as the roller/squeegee move across the surface.

The application of a target solution to microwell arrays results in the deposition of the solution at each of the microwell locations. The chemical and/or mechanical barriers would maintain the integrity of the array and prevent cross-contamination of reagents from element to element. The reagents printed at each microwell would be rehydrated by the target solution resulting in an ultra-low volume reaction mix. In some embodiments, the microwell-microarray reactions are covered with mineral oil or some other suitable evaporation barrier to allow high temperature incubation. The signal generated may be detected directly through the applied evaporation barrier using a fluorescence microscope, array reader or standard fluorescence plate reader.

Advantages of the use of a microwell-microarray, for running INVADER assays (e.g. dried down INVADER assay components in each well) include, but are not limited to: the ability to use the INVADER Squared (Biplex) format for a DNA detection assay; sufficient sensitivity to detect hgDNA directly, the ability to use "universal" FRET cassettes; no attachment chemistry needed (which means already existing off the shelf reagents could be used to print the microarrays), no need to fractionate hgDNA to account for surface effect on hybridization, low mass of hgDNA needed to make tens of thousands of calls, low volume need (e.g. a 100μm microwell would have a volume of 0.28nl, and at 10^4 microwells per array a volume of 2.8μl would fill all wells), a solution of 333ng/μl hgDNA would result in ~100 copies per microwell (this is 33X more concentrated than the use of 100ng hgDNA in a 20μl reaction), thus 2.8μl x 333ng/μl = 670 ng hgDNA for 10^4 calls or 0.07 ng per call. It is appreciated that other detection assays can also be presented in this format.

C. Distribution, Use, and Pricing of Detection Assays

As discussed above, the use of detection assays in the context of research products using the systems and methods of the present invention generates data (which can in one variant be sent automatically over a computer network to one or more components of the present invention)

that finds use in obtaining regulatory approval for clinical products and in the generation of databases, which also optionally are used with components of the present invention. In some embodiments of the present invention, a party with interest in selling products (e.g., clinical products) or information stored in databases provides (e.g., using any delivery systems) detection assays to researchers in order to collect data. In some embodiments, the party provides detection assays to researchers at a reduced cost, at a subsidized cost, or at no cost in order to receive data from said researchers. In yet other embodiments, the party pays a researcher to use the test in order to gain access to data obtained from the test for use in the components hereof. Using the systems and methods of the present invention, the party can compensate for any lost profits or revenues by obtaining and selling clinical products, which are typically high revenue, high margin products.

In one variant of the invention, the system and method of the present invention includes a consumer direct web order entry component (see above). The consumer direct web order entry component provides one or more interactive screens or web pages on a consumer's computer, which is accessible over the Internet or other computer network, from which a consumer can order oligonucleotide detection assay services to be conducted on a genetic sample of the consumer. The consumer can directly order detection assays of the consumer's genetic material or precursor material, e.g. whole blood or other material, through these interactive screens or web pages. In one variant of the invention, the customer can search by allelle frequency. The web pages present the consumer with various assays, panels of detection assays, e.g. a DME panel or screen, or a cardiovascular panel or screen, assays from different manufacturers, and/or combinations thereof. The consumer chooses which detection assay or panel of detection assays the consumer would like to order. The consumer inputs his data on the web page or screen, including but not limited to name, address information, credit card information or other billing or payment information, detection assay, screen or panel selection information from a plurality of different options. This information is then sent to a host computer or server. The host computer or server processes this information and sends the consumer a kit for taking a sample of the consumer's genetic material, e.g. whole blood via a pin prick and collection container, with appropriate identifying markings linking the kit to the consumer and the requisite detection assays or panel(s) requested. The consumer sends back kit with the genetic material or precusor

material back to a service provider which then correlates the sample shipment to a predetermined detection assay or panel product, processes the sample, analyzes the sample, and sends the results back to the consumer via the web, e.g. using e-mail, or via a report sent by standard mail. In one alternative of the invention, the consumer logs back on to the web order entry component to access his or her result data by entering a password provided to the consumer upon placement of the intial order or at some latter time.

It is appreciated that this approach provides the consumer with access to personalized medical information, and increases the amount and timeliness of information the consumer is provided with so that informed medical decisions can be made. It is appreciated that the consumer can also have access to an on-line Physician's Desk Reference ("PDR") (which may be located on the same or different site from that of the consumer direct web order entry component) which has drug information correlated with detection assay information. The Physician's Desk Reference is incorporated herein by reference as if fully set forth. By way of further example, a consumer may be taking a drug which may not be effective to treat the consumer's medical condition. The consumer logs onto the consumer direct web order entry component and enters the name of his drug. He is provided with PDR drug information correlated to detection assay information, e.g. the type of detection assay or panel that should be provided when deciding whether or not to use or prescribe the drug. The consumer then orders the detection assay or detection panel screening service as described above from the service provider, and receives the results of the screen. The results indicate that the consumer has a DME profile such that the drug originally given to the consumer would not be effective or have reduced effectiveness. The consumer is then provided with drug alternatives that are effective for consumer's with this genetic profile. The patient can then approach his physician with this information and seek a prescription for the other drug altenatives and discontinue use of the ineffective drug. It is appreciated that this system and method can also be used proactively prior to the presciption of a drug or drug combination therapy to select the best drug or combination of drugs depending on the consumer's genetic profile. In this variant of the invention, it is appreciated that the PDR is in an electronic format and individual drug entries of information are correlated with data of one or more detection assays or detection assay panel data. In one variant of the invention the PDR forms an integral part of the web order entry component of the

invention. In yet another variant, the invention provides a link to the electronic PDR which may be located on another web site.

It is appreciated that the customer order entry component and/or the billing component comprise, in one variant of the invention, a differential pricing component. The differential pricing component is a routine or set of routines that run on one or more computers or other circuitry of the system that provide the ability to price detection assays by the category of detection assay purchased by the consumer or other entity. The billing component may include a secure web based transaction billing routine or software packages, or standard billing routines or software packages commercially available providing billing and tracking functionality. It is also appreciated that the detection assay locator component is periodically update with additional detection assays that are available and are offered for sale.

By way of example, detection assay A or detection assay panel B is either an RUO product, an ASR product, or an IVD product. It is appreciated that in one version of the invention there is subtantially no difference or no difference between and RUO product, an ASR product, or an IVD product except for price and/or the quality control process the detection assay undergoes, if any. In some embodiments, there is differential pricing for 1) new products (e.g. assays that have not been designed or produced before), 2) low volume products, 3) high volume products, 4) single components of an assay, and 5) an entire kit. In one version of the invention, a customer selects detection assay A or detection assay panel B. The web page then displays a choice between detection assay A-RUO product, detection assay A-ASR product, detection assay B-IVD product. The consumer selects which type of product he desires, e.g. RUO product. The selection is then sent to the remote host computer, and a corresponding RUO product price is presented to the consumer. In another variant, the consumer chooses detection assay A-IVD product. Upon selecting this option the user is display a different price, e.g. an IVD product price. The transaction is then processed. It is also appreciated that that billing component also makes use of this differential pricing feature so that records of the transactions are processed properly. In further embodiments, systems of the present invention also indicate if their is intellectual property (IP) that may cause the prive of the detection assay to increase (e.g. detection assay provided may have paid for a license already, may need to pay a license fee, or may be risking patent litigation through the sale of the assay).

It is also appreciated that the differential pricing routines are capable of pricing the detection assay based upon the platform that the customer selects for the single detection assay or a plurality of detection assays. For example, if a customer selects a 96 well format, price data A are correlated to the detection assay and the transaction is processed. If the customer selects a 384 well format, price data B are correlated to the detection assay and the customer total is appropriately calculated.

D. Medical Records

The present invention also relates to medical records (e.g., electronic medical records) comprising genetic information (e.g., patient-specific genetic information) obtained from using one or more of the detection assays produced by the systems and methods described herein. In particular the present invention provides systems and methods for the generation of large amounts of genetic information related to medically relevant conditions and the use of this information in patient health care. For example, the present invention provides systems and methods for generating clinically valid polymorphism data (e.g., SNP data) for any desired subject or population. The data includes information about the presence or absence of the polymorphism in a test subject and a correlation between the presence of a polymorphism or set of polymorphisms and one or more medically relevant conditions. In one variant, this information is generated at a plurality of remote nodes at detection assay user sites and then communicated to one or more central nodes for processing thereof. This information finds use in many aspects of patient health care, including, but not limited to, selection of prescriptions, avoidance of undesired drug reactions or allergic reactions, selection of medical courses of action or therapeutic routes, and the like. Therefore, this information forms a valuable part of the patient's medical records for use in nearly every aspect of patient care. As such, the present invention provides medical records electronically that contain useful genetic information as well as other patient data including, but not limited to prescription data (e.g., data related to one or more drugs or other prescribed medical interventions of the subject, including drug identity, drug reaction data, allergies, risk assessment data, and multi-drug interaction data, billing code levels, order restrictions); information pertaining a physician visit (e.g., date and time of visit, identity of physicians, physician notes, diagnosis information, differential diagnosis information, patient

location, patient status, order status, referral information); patient identification information (e.g., patient age, gender, race, insurance carrier, allergies, past medical history, family history, social history, religion, employer, guarantor, address, contact information, patient condition code); and laboratory information (e.g., labs, radiology, and tests).

The genetic information of the present invention may be incorporated into any type of medical record system including electronic medical record systems (e.g., U.S. Pat. Nos. 6,272,468, 6,266,645, 6,263,330, 6,246,975, 6,234,964, 6,206,829, 6,192,112, 6,113,540, 6,088,677, 6,071,236, 6,022,315, 6,006,191, 5,974,398, 5,950,168, 5,924,074, 5,910,107, 5,890,129, 5,867,821, 5,845,255, 5,832,450, 5,823,948, 5,737,539, and PCT Publication Nos. WO 01/54571, WO 00/28460, WO 00/65522, WO 00/29983, WO 00/28459, and WO 99/21114, each of which is herein incorporated by reference in its entirety.

The present invention is not limited by the process of incorporating genetic information into medical records. In some embodiments, genetic information is added to pre-existing medical records, and the data correlated thereto. For example, a subjects electronic medical record is stored on a computer system of a health care professional or an agency that houses data for health care professionals. The genetic information is received by the computer system and stored as part of the medical record. In some embodiments, the genetic information is manually entered into the electronic medical record. In other embodiments, the genetic information is transmitted to the computer system housing the medical record using a communications network (e.g., the Internet). For example, in some embodiments, genetic information (e.g., polymorphism information) is directly transmitted over a communications network from a computer system designed to collect and/or store the genetic information to the computer system housing the medical record. In some embodiments, genetic information is used to create an electronic medical record, wherein additional information pertaining to the subject is added along with, or subsequently, to the medical record.

Genetic information contained in a medical record of the present invention is retrieved and used at any desired time by any desired party. Genetic information, alone, or in combination with other information contained in the medical record, finds use in selecting appropriate health care decisions and courses of action. The health care professional, or other users, evaluate the genetic information, along with other information about the subject in making a informed

decision based on all of the circumstances and using the individual's profession judgment. For example, a physician, upon viewing the genetic information and other information contained in the medical record may elect to schedule a medical procedure. Likewise, a pharmacy may elect to prepare a particular type of medication or dose of medication or avoid certain medications based on the information contained in the medical record.

In some embodiments, genetic information is linked to preexisting medical records to enhance the analysis of the genetic information. For example, in some embodiments, a plurality (e.g., thousands) of patient samples are tested to determine one or more genetic characteristics. This genetic information is then compared with the patient's preexisting medical records to determine correlations between the genetic identity and one or more characteristics of the patient contained in the medical record. This allows genetic information (e.g., SNPs) to be correlated to particular medical conditions, drug interactions, gender, race, or other patient characteristics.

In some embodiments of the present information, genetic information contained in a medical record is derived from a biological detection assay, including an indication of the presence or absence of a polymorphism in a subject that is correlated with a medically relevant condition. The present invention is not limited by the identity of the detection assay. For example, in some preferred embodiments, the detection assay is an invasive cleavage assay (e.g., the INVADER assay, Third Wave Technologies, Madison, WI) or other detection assay described herein. The present invention provides tens of thousands of designed detection assays (e.g., the INVADER detection assays provided in Figure 6). The detection assays in Figure 6 or equivalent assays (e.g., assays targeting similar target sequences, assays using similar probe sequences, non-invasive cleavage assays that use one or more component shown in Figure 6 or designed based on one or more components shown in Figure 6, e.g., other hybridization methods using one or more sequences similar to those in Figure 6) are used to generate genetic information. In other preferred embodiments, other detection assay technologies are used to generate genetic information for use in the medical records of the present invention.

All publications and patents mentioned in the above specification are herein incorporated by reference as if expressly set forth herein. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described

in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in relevant fields are intended to be within the scope of the following claims.

CLAIMS

WE CLAIM:

1. A system for manufacturing and selling detection assays, comprising:
   a. a computer-based customer order component for ordering at least one of a plurality of oligonucleotide detection assays;
   b. a detection assay production component for creating said oligonucleotide detection assays;
   c. a shipping component for shipping said oligonucleotide detection assays; and
   d. a billing component for billing a customer for said oligonucleotide detection assays.

2. The system of Claim 1, wherein said billing component comprises a payment receipt component for receiving payment for said oligonucleotide detection assays.

3. The system of Claim 1, wherein said computer-based customer order component comprises a client-based computer network.

4. The system of Claim 1, wherein said computer-based customer order component comprises a distributor-based computer network.

5. The system of Claim 1, wherein said computer-based customer order component comprises a web-based user interface for ordering said oligonucleotide detection assay.

6. The system of Claim 5, wherein said web-based user interface provides a detection assay locator component.

7. The system of Claim 6, wherein said detection assay locator component comprises a library of detection assay data from which said oligonucleotide detection assay can be selected.

8. The system of Claim 7, wherein said library of detection assay data comprising single nucleotide polymorphism data or RNA data.

9. The system of Claim 1, wherein said detection assay production component comprises a shop floor control system.

10. The system of Claim 9, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a make-to-order routine.

11. The system of Claim 9, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a make-to-stock routine.

12. The system of Claim 9, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a fulfill-from-stock routine.

13. The system of Claim 9, wherein said shop floor control system comprises a library of detection assay data from which said plurality of detection assay can be created.

14. The system of Claim 1, wherein said detection assay production component comprises a synthesis component.

15. The system of Claim 1, wherein said detection assay production component comprises a cleave/deprotect component.

16. The system of Claim 1, wherein said detection assay production component comprises a purification component.

17. The system of Claim 1, wherein said detection assay production component comprises a dilute and fill component.

18. The system of Claim 1, wherein said detection assay production component comprises a quality control component.

19. The system of Claim 14, wherein said synthesis component comprises a plurality of oligonucleotide synthesizers.

20. The system of Claim 19, wherein said plurality of oligonucleotide synthesizers are selected from the group consisting of MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), OligoPilot (Amersham Pharmacia,), the 3900 and 3948 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA), POLYPLEX (Genemachines), 8909 EXPEDITE, Blue Hedgehog (Metabio), MerMade (BioAutomation, Plano, Texas), Polygen (Distribio, France), and PrimerStation 960 (Intelligent Bio-Instruments, Cambridge, MA).

21. The system of Claim 1, wherein said detection assay production component comprises an inventory control component.

22. The system of Claim 1, wherein said oligonucleotide detection assay comprises an invasive cleavage assay.

23. The system of Claim 1, wherein said oligonucleotide detection assay comprises a TAQMAN assay.

24. The system of Claim 1, wherein said oligonucleotide detection assay comprises an assay selected from the group consisting of a sequencing assay, a polymerase chain reaction assay, a hybridization assay, a hybridization assay employing a probe complementary to a

mutation, a microarray assay, a bead array assay, a primer extension assay, an enzyme mismatch cleavage assay, a branched hybridization assay, a rolling circle replication assay, a NASBA assay, a molecular beacon assay, a cycling probe assay, a ligase chain reaction assay, and a sandwich hybridization assay.

25. The system of Claim 1, wherein said oligonucleotide detection assay is configured to detect a sequence selected from the group consisting of a polymorphism, a transgene, a splice junction, a mammalian sequence, a prokaryotic sequence, and a plant sequence.

26. The system of Claim 1, wherein said detection assay production component comprises an oligonucleotide detection assay design component.

27. The system of Claim 26, where said detection assay design component comprises a PCR primer creation component.

28. The system of claim 27, wherein said PCR primer creation component is configured to optimizer PCR primer concentrations.

29. The system of Claim 26, wherein said detection assay design component is configured to design a plurality of detections assays for detecting the presence of one or more polymorphisms.

30. The system of Claim 1, wherein said order entry component or said billing component comprises a differential pricing component.

31. The system of Claim 30, wherein said differential pricing component is capable of selectably pricing said detection assay based upon a predetermined category of product.

32. The system of Claim 31, wherein said predetermined category of product is selected from the group consisting of an RUO product, an ASR product, and an IVD product.

33. The system of Claim 30, wherein said differential pricing component comprises a routine that associates a predetermined price of a detection assay based upon a presentation platform selection.

34. The system of Claim 1 in which said computer based customer order entry component further comprises a consumer direct web order entry component.

35. The system of Claim 1 in which said computer-based customer order entry component provides a data feed into said detection assay production component.

36. The system of Claim 35 in which said data feed affects production of said oligonucleotide detection assays.

37. The system of Claim 35 in which said data feed comprises statistical information associated with one or more oligonucleotide detection assays.

38. The system of Claim 37 in which said statistical information is selected from the group consisting of total oligonucleotide detection assays ordered or oligonucleotide detection assay orders received; a histogram; an oligonucleotide detection assay average per consumer; an arithemetic mean; quantity of oligonucleotide detection assays, size of order of oligonucleotide detection assays; format of panel information; a mode; a median; a weighted mean; a harmonic mean; a geometric mean; a logarithmic mean; a root mean square; a root sum square, and combination thereof; a normal distribution curve, said normal distribution curve selected from the group consisting of a normal distribution curve of number of consumers, number of detection assays, quantity of oligonucleotide detection assays, quantity of oligonucleotide detection assays or a certain type; a spread; a variance; a standard deviation; a skewed distribution; a sampling; a confidence level; and, a regression analysis.

39. A detection assay ordering system, comprising a first processor in electronic communication with: a) a computer system of a customer; b) a second processor configured to carry out detection assay design; and c) a third processor configured to carry out detection assay production.

40. The system of Claim 39, wherein said detection assay comprises an invasive cleavage assay.

41. The system of Claim 39, wherein said first processor provides a user interface to said computer system of said customer.

42. The system of Claim 41, wherein said user interface comprises stacked databases.

43. The system of Claim 42, wherein said stacked databases comprise SNP data.

44. The system of Claim 42, wherein said stacked databases comprise pre-existing detection assay data.

45. The system of Claim 44, wherein said pre-existing detection assay comprises a detection assay that has passed through an in silico process.

46. The system of Claim 44, wherein said pre-existing detection assay comprises a detection assay that has passed through a genotyping process.

47. A method for screening candidate oligonucleotides for use in a detection assay, comprising:

a) providing

i) a candidate oligonucleotide;

ii) five or more target nucleic acids, wherein each of said five or more target nucleic acids is derived from a different subject; and

iii) detection assay components that permit detection of said target nucleic acids in the presence of a functional detection oligonucleotide; and

b) treating together said five or more target nucleic acids with said candidate oligonucleotide in the presence of said detection assay components;

c) determining if said candidate oligonucleotide is a functional detection oligonucleotide for use with each of said five or more target nucleic acids.

48. The method of Claim 47, wherein said five or more target nucleic acids comprises 50 or more target nucleic acids derived from different subjects.

49. The method of Claim 48, wherein said 50 or more target nucleic acids comprises 100 or more target nucleic acids derived from different subjects.

50. The method of Claim 47, wherein one or more of said target nucleic acids comprises a single nucleotide polymorphism.

51. The method of Claim 47, wherein said candidate oligonucleotide comprises a hybridization probe.

52. The method of Claim 51, wherein said candidate oligonucleotide is designed to hybridize to a target sequence of at least one of said target nucleic acids.

53. The method of Claim 52, wherein said target sequence is identified by in silico analysis.

54. The method of Claim 47, wherein said detection assay components comprise detections assay components for performing an INVADER assay.

55. The method of Claim 47, further comprising the step of d) preparing a kit containing said candidate oligonucleotide if said candidate oligonucleotide is determined to be a functional detection oligonucleotide.

56. The method of Claim 55, wherein said kit comprises instructions, directing a user of said kit to use said kit with samples from subjects suspected of possessing any of said target nucleic acids from which said candidate oligonucleotide was determined to be a functional detection oligonucleotide.

57. A method of gathering and storing genomic data derived from a detection assay, comprising:

a) providing:
   i) a detection assay configured to detect the presence or absence of a nucleic acid sequence in a sample;
   ii) a first computer system comprising one or more computer processors and a computer memory;
   iii) a second computer system comprising one or more computer processors and computer memory, wherein said computer memory comprises a genomic information database; and
   iv) a test sample
b) treating said test sample with said detection assay to generate test result data;
c) collecting said test result data with said first computer system;
d) transmitting said test result data from said first computer system to said second computer system under conditions such that said test result data is added to said genomic information database of said second computer system.

58. The method of Claim 57, wherein said detection assay comprises an assay selected from the group consisting of hybridization assays, cleavage assays, amplification assays, sequencing assays, and ligation assays.

59. The method of Claim 57, wherein said detection assay comprises an assay selected from the group consisting of INVADER assays and TAQMAN assays.

60. The method of Claim 57, wherein said nucleic acid sequence comprises a single nucleotide polymorphism.

61. The method of Claim 57, wherein said first computer systems comprises a detector.

62. The method of Claim 57, wherein said detector is selected from the group consisting of fluorescence detectors, luminescence detectors, optical detectors, and radioactivity detectors.

63. The method of Claim 57, wherein said test sample comprises a genomic DNA sample.

64. The method of Claim 57, wherein said test sample comprises an RNA sample.

65. The method of Claim 57, wherein said test result data comprise information related to a subject from which said test sample was derived.

66. The method of Claim 57, wherein said first computer system is located in a different geographic location from said second computer system.

67. The method of Claim 57, wherein said transmitting comprising sending said test result data over a communication network.

68. The method of Claim 57, wherein said test result data comprises allele frequency information.

69. The method of Claim 57, wherein said genomic information database comprises database data comprising allele frequency information.

70. A method for searching nucleic acid databases comprising:

a) providing

i) a central node comprising a processor;

ii) a plurality of sub-nodes in electronic communication with said central node, said sub-nodes comprising sequence database information; and

iii) nucleic acid sequence to be searched;

b) providing said nucleic acid sequence to be searched to said central node;

c) concurrently sending said nucleic acid sequence information to be searched from said central node to said plurality of sub-nodes;

d) searching said sequence database information with said nucleic acid sequence to be searched to generate search results.

71. The method of Claim 70, further comprising the step of e) sending said search results from said plurality of sub-nodes to said central node.

72. The method of Claim 71, wherein steps b), c), d), and e) are complete in two seconds or less.

73. The method of Claim 71, wherein two or more distinct sequence databases are stored on said plurality of sub-nodes.

74. The method of Claim 73, wherein one of said two or more distinct sequence databases is stored on two or more of said plurality of sub-nodes.

75. The method of Claim 73, wherein two or more copies of said two or more distinct sequence databases are stored on said plurality of sub-nodes.

76. The method of Claim 71, wherein each of said plurality of sub-nodes comprises a single sequence database.

77. The method of Claim 71, wherein said nucleic acid sequence to be searched comprises a single nucleotide polymorphism.

78. The method of Claim 71, wherein said sequence database information comprises one or more databases selected from the group consisting of GoldenPath, GenBank, dbSNP, UniGene, LocusLink, and SNP Contorium databases.

79. A method for characterizing a target sequence comprising:

a. screening said target sequence for the presence of repeat sequences and heterologous sequences to generate a masked target sequence;
b. searching a plurality of sequence databases with said masked target sequence to generate search result data; and
c. generating a report comprising said search result data.

80. The method of Claim 79, wherein said plurality of sequence databases comprises one or more databases selected from the group consisting of polymorphism databases, genome databases, linkage databases, and disease association databases.

81. The method of Claim 79, wherein said plurality of sequence databases comprises one or more databases selected from the group consisting of GoldenPath, GenBank, dbSNP, UniGene, LocusLink, and SNP Contorium databases.

82. The method of Claim 79, wherein said target sequence comprises a single nucleotide polymorphism.

83. The method of Claim 79, wherein said report provides a reliability score, said reliability score representing a likelihood of success of detecting said target sequence performance in a detection assay.

84. The method of Claim 79, wherein said report indicates the presence or absence of said target sequence in one or more of said plurality of sequence databases.

85. The method of Claim 79, wherein said report indicates a position of said target sequence in a genome.

86. The method of Claim 79, wherein said report provides polymorphism information related to said target sequence.

87. A database comprising allele frequency information, said allele frequency information generated by a method comprising:

a) producing a detection assay for detecting a target sequence;
b) testing five or more target sequences from different subjects with said detection assay to produce assay data;
c) storing said assay data in a database, wherein said assay data is correlated to at least one characteristic of said subjects.

88. The database of Claim 87, wherein said target sequence comprises a single nucleotide polymorphism.

89. The database of Claim 87, wherein said five or more target sequences comprise 50 or more target sequences from different subjects.

90. The database of Claim 87, wherein said five or more target sequences comprise 100 or more target sequences form different subjects.

91. The database of Claim 87, wherein said at least one characteristic of said subjects comprises subject race.

92. The database of Claim 87, wherein said at least one characteristic of said subjects comprises presence of a disease state.

93. A method for collecting genomic information comprising:

a) providing:

i) a detection assay that detects the presence of a target nucleic acid sequence in a sample;

ii) a software application on a computer system of a user, said software application configured to receive detection assay data;

iii) a database on a computer system of a service provider;

iv) a communications network; and

v) one or more samples comprising nucleic acid;

b) treating said one or more samples with said detection assay to generate assay data;

c) collecting said assay data with said software application;

d) transmitting said assay data from said computer system of said user to said computer system of said service provider using said communications network; and

e) storing said assay data in said database.

94. The method of Claim 93, wherein said target nucleic acid sequence comprises a single nucleotide polymorphism and wherein said detection assay detects the presence or absence of said single nucleotide polymorphism.

95. A database generated by the method of Claim 93.

96. A method of detecting single nucleotide polymorphisms comprising:

a) providing:

i) a plurality of samples comprising genomic DNA from a first individual and four or more additional individuals, each of said individuals having genomic DNA comprising a first region, said first individual having a first single nucleotide polymorphism in said first region;

ii) at least one detection reagent capable of generating a signal; and

iii) at least one oligonucleotide probe designed to cause said detection reagent to generate a signal following contact of said probe with a portion of said first region of said genomic DNA of said first individual;

b) contacting each of said genomic DNA samples with said oligonucleotide probe under conditions such that a signal is detected for said genomic DNA of said first individual;

c) detecting at least one of said four or more additional individuals for which no signal is detected, thereby detecting a negative-tested individual; and

d) assaying said first region of said negative-tested individual under conditions such that a second single nucleotide polymorphism is revealed in said first region of said genomic DNA of said negative-tested individual in addition to said first single nucleotide polymorphism, wherein said first individual lacks said second single nucleotide polymorphism.

97. The method of Claim 96, further providing a second oligonucleotide probe designed to cause said detection reagent to generate a signal following contact of said probe with a portion of said first region of said genomic DNA of said negative-tested individual, wherein said second oligonucleotide probes is contacted with said genomic DNA sample of said negative-tested individual.

98. A method comprising;

a) providing target sequence information for at least Y target sequences, wherein each of said target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of said footprint region, and iii) a 3' region immediately downstream of said footprint region, and

b) processing said target sequence information such that a primer set is generated, wherein said primer set comprises a forward and a reverse primer sequence for each of said at least Y target sequences, wherein each of said forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of said forward and reverse primers is not complementary to N[2]-N[1]-3' of any of said forward and reverse primers in said primer set.

99. The method of Claim 98, wherein said primer set is configured for performing a multiplex PCR reaction that amplifies at least Y amplicons, wherein each of said amplicons is defined by the position of said forward and reverse primers.

100. The method of Claim 98, wherein said primer set is generated as digital or printed sequence information.

101. The method of Claim 98, wherein said primer set is generated as physical primer oligonucleotides.

102. The method of Claim 98, wherein N[3]-N[2]-N[1]-3' of each of said forward and reverse primers is not complementary to N[3]-N[2]-N[1]-3' of any of said forward and reverse primers in said primer set.

103. The method of Claim 98, wherein said processing comprises initially selecting N[1] for each of said forward primers as the most 3' A or C in said 5' region.

104. The method of Claim 98, wherein said processing comprises initially selecting N[1] for each of said reverse primers as the most 3' A or C in the complement of said 3' region.

105. The method of Claim 98, wherein said footprint region comprises a single nucleotide polymorphism.

106. The method of Claim 8, wherein said footprint region for each of said target sequences comprises a portion of said target sequence that hybridizes to one or more assay probes configured to detect said single nucleotide polymorphism

107. The method of Claim 98, wherein said processing further comprises selecting N[5]-N[4]-N[3]-N[2]-N[1]-3' for each of said forward and reverse primers such that less than 80 percent homology with a assay component sequence is present.

108. The method of Claim 98, wherein Y is an integer between 2 and 500.

109. The method of Claim 98, wherein said processing comprises selecting x for each of said forward and reverse primers such that each of said forward and reverse primers has a melting temperature with respect to said target sequence of approximately 50 degrees Celsius.

110. The method of Claim 98, wherein forward and reverse primer pair optimized concentrations are determined for said primer set.

111. The primer set generated by the method of Claim 98.

112. A method comprising;

a) providing;

i) a user interface configured to receive sequence data,

ii) a computer system having stored therein a multiplex PCR primer software application, and

b) transmitting said sequence data from said user interface to said computer system, wherein said sequence data comprises target sequence information for at least Y target sequences, wherein each of said target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of said footprint region, and iii) a 3' region immediately downstream of said footprint region, and

c) processing said target sequence information with said multiplex PCR primer pair software application to generate a primer set, wherein said primer set comprises; i) a forward primer sequence identical to at least a portion of said target sequence immediately 5' of said footprint region for each of said Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of said target sequence immediately 3' of said footprint region for each of said at least Y target sequences, wherein each of said forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of said forward and reverse primers is not complementary to N[2]-N[1]-3' of any of said forward and reverse primers in said primer set.

113. The method of Claim 112, wherein said primer set is configured for performing a multiplex PCR reaction that amplifies at least Y amplicons, wherein each of said amplicons is defined by the position of said forward and reverse primers.

114. The method of Claim 112, wherein said primer set that is generated is digital or printed sequence information.

115. The method of Claim 112, wherein N[3]-N[2]-N[1]-3' of each of said forward and reverse primers is not complementary to N[3]-N[2]-N[1]-3' of any of said forward and reverse primers in said primer set.

116. The method of Claim 112, wherein said processing comprises initially selecting N[1] for each of said forward primers as the most 3' A or C in said 5' region.

117. The method of Claim 112, wherein said processing comprises initially selecting

118. The method of Claim 112, wherein said footprint region comprises a single nucleotide polymorphism.

119. The method of Claim 118, wherein said footprint region for each of said target sequences comprises a portion of said target sequence that hybridizes to one or more assay probes configures to detect said single nucleotide polymorphism.

120. The method of Claim 112, wherein Y is a number between 2 and 500.

121. The method of Claim 112, wherein said processing comprises selecting x for each of said forward and reverse primers such that each of said forward and reverse primers has a melting temperature of approximately 50 degrees Celsius.

122. The primer set generated by the method of Claim 112.

123. A system comprising;

a) a computer system configured to receive data from a user interface, wherein said user interface is configured to receive sequence data, wherein said sequence data comprises target sequence information for at least Y target sequences, wherein each of said target sequences comprises; i) a footprint region, ii) a 5' region immediately upstream of said footprint region, and iii) a 3' region immediately downstream of said footprint region,

b) a multiplex PCR primer pair software application operably linked to said user interface, wherein said multiplex PCR primer software application is configured to process said target sequence information to generate a primer set, wherein said primer set comprises; i) a forward primer sequence identical to at least a portion of said target sequence immediately 5' of said footprint region for each of said Y target sequences, and ii) a reverse primer sequence identical to at least a portion of a complementary sequence of said target sequence immediately 3' of said footprint region for each of said at least Y target sequences, wherein each of said forward and reverse primer sequences comprises a nucleic acid sequence represented by 5'-N[x]-N[x-1]- ....-N[4]-N[3]-N[2]-N[1]-3', wherein N represents a nucleotide base, x is at least 6, N[1] is nucleotide A or C, and N[2]-N[1]-3' of each of said forward and reverse primers is not complementary to N[2]-N[1]-3' of any of said forward and reverse primers in said primer set, and

c) a computer system having stored therein said multiplex PCR primer pair software application, wherein said computer system comprises computer memory and a computer processor.

124. The system of Claim 123, wherein said computer system is configured to return said primer set to said user interface.

125. The system of Claim 123, wherein said multiplex PCR primer software application is further configured to process said target sequence information such that x is selected for each of said forward and reverse primers such that each of said forward and reverse primers has a melting temperature of approximately 50 degrees Celsius.

126. An nucleic acid synthesis reagent delivery system comprising:

a. one or more reagent containers containing nucleic acid synthesis reagent;
b. a branched delivery component attached to said one or more reagent containers such that said nucleic acid synthesis reagent can pass from said reagent containers to said branched delivery component, wherein said branched delivery component comprises a plurality of branches;

c. a plurality of delivery lines, said plurality of delivery lines attached on one end to a branch of said branched delivery component and attached on a second end to a nucleic acid synthesizer.

127. The system of Claim 126, wherein said plurality of branches comprises ten or more branches.

128. The system of Claim 126, wherein said plurality of delivery lines comprises ten or more delivery lines.

129. The system of Claim 126, wherein said branched delivery component comprises a sight glass.

130. The system of Claim 129, wherein said sight glass comprises a purge valve.

131. The system of Claim 126, wherein one or more of said plurality of delivery lines comprises a shut-off valve.

132. A waste disposal system comprising:

a. a waste tank comprising a waste input channel configured to receive liquid waste product and a waste output channel configured to remove liquid waste when said waste tank is purged; and

b. a pressurized gas line attached to said waste tank, said pressurized gas line configured to deliver gas into said waste tank when said waste tank is to be purged, wherein said gas line is configured to deliver a gas that allows purging of said waste tank.

133. The system of Claim 132, wherein said pressurized gas line is attached to an argon gas source.

134. The system of Claim 132, wherein said gas is delivered at a low pressure.

135. The system of Claim 134, wherein said low pressure is 10 pounds per square inch or less.

136. The system of Claim 135, wherein said low pressure is 5 pounds per square inch or less.

137. The system of Claim 132, wherein said waste input channel is attached to a waste line, said waste line attached to a plurality of nucleic acid synthesizers.

138. The system of Claim 134, wherein said plurality of nucleic acid synthesizers comprises twenty or more nucleic acid synthesizer.

139. The system of Claim 132, wherein said waste tank further comprises a sight glass.

140. The system of Claim 132, further comprising an automated purge component, said automated purge component capable of detecting waste levels in said waste tank and purging said waste tank when said waste levels are at or above a threshold level.

141. A method for purifying nucleic acids comprising:

a. providing:
    i. an nucleic acid purification column;
    ii. a buffer;
    iii. a nucleic acid mixture;
b. contacting said nucleic acid mixture with said nucleic acid purification column;
c. adding said buffer to said nucleic acid purification column, wherein a nucleic acid molecule having between 23-39 nucleotides is eluted from said nucleic acid purification column in less than forty minutes.

142. The method of Claim 141, wherein said nucleic acid purification column is contained in an HPLC apparatus.

143. A method for deprotecting nucleic acid molecules comprising:

a. providing:
   i. a multiwell plate configured to hold a plurality of protected nucleic acid molecules;
   ii. a plurality of different protected nucleic acid molecules;
b. placing said nucleic acid molecules into said multiwell plates; and
c. treating said plate under conditions that resulted in the deprotection of said nucleic acid molecules.

144. The method of Claim 143, wherein said multiwell plate comprises a 96-well plate.

145. A nucleic acid synthesizer comprising a plurality of synthesis columns and an energy input component that imparts energy to said plurality of synthesis columns to increase nucleic acid synthesis reaction rate in said plurality of synthesis columns.

146. The synthesizer of Claim 145, further comprising a fail-safe reagent delivery component configured to deliver one or more reagent solutions to said plurality of synthesis columns.

147. The synthesizer of Claim 146, wherein said fail-safe reagent delivery component comprises a plurality of reagent tanks.

148. The synthesizer of Claim 147, wherein said plurality of reagent tanks comprise one or more tanks selected from the group consisting of acetonitrile tanks, phosphoramidite tanks, argon gas tanks, oxidizer tanks, tetrazole tanks, and capping solution tanks.

149. The synthesizer of Claim 148, wherein said reagent tanks further comprise a plurality of large volume containers, each said large volume container comprising at least one of said reagent solutions.

150. The synthesizer of Claim 149, wherein said large volume containers store in the range of about 2 liters to about 200 liters of said one or more reagent solutions.

151. The synthesizer of Claim 145, wherein said energy input component comprises a heating component.

152. The synthesizer of Claim 151, wherein said heating component provides substantially uniform heat to said plurality of synthesis columns.

153. The synthesizer of Claim 145, wherein said energy input component provides heated reagent solutions to said plurality of synthesis columns.

154. The synthesizer of Claim 145, wherein said energy input heats said plurality of synthesis columns in the range of about 20 to about 60 degrees Celsius.

155. The synthesizer of Claim 145, wherein said energy input component comprises a heating coil.

156. The synthesizer of Claim 145, wherein said energy input component comprises a heat blanket.

157. The synthesizer of Claim 145, wherein said energy input component comprises a heated room.

158. The synthesizer of Claim 145, wherein said energy input component provides energy in the electromagnetic spectrum.

159. The synthesizer of Claim 145, wherein said energy input component comprises an oscillating member.

160. The synthesizer of Claim 145, wherein said energy input component provides a periodic energy input.

161. The synthesizer of Claim 145, wherein said energy input component provides a constant energy input.

162. The synthesizer of Claim 151, wherein said heating component comprises a resistance heater.

163. The synthesizer of Claim 151, wherein said heating component comprises a Peltier device.

164. The synthesizer of Claim 151, wherein said heating component comprises a magnetic induction device.

165. The synthesizer of Claim 151, wherein said heating component comprises a microwave device.

166. The synthesizer of Claim 151, wherein said heating component comprises heated fluid or gas..

167. The synthesizer of Claim 145, further comprising a mixing component that mixes reagents in said plurality of synthesis columns.

168. The synthesizer of Claim 167, wherein said mixing component is selected from the group consisting of an ultrasonic mixer, a magnetic mixer, a fluid oscillator, and a vibrational mixer.

169. The synthesizer of Claim 145, further comprising a reaction support, said reaction support configured to hold three or more synthesis columns.

170. The synthesizer of Claim 169, wherein said reaction support is configured for operation with a cleavage and deprotect component.

171. The synthesizer of Claim 170, further comprising sample tracking software configured to associated sample identification tags with samples that are processed by said synthesizer and said cleavage and deprotect component.

172. The synthesizer of Claim 171, wherein said sample tracking software is further configured to receive synthesis request information from a user, prior to sample processing by said synthesizer.

173. The synthesizer of Claim 170, further comprising a robotic component configured to transfer said reaction support from said synthesizer to said cleavage and deprotect component.

174. The synthesizer of Claim 173, wherein said robotic component is further configured to transfer said reaction support from said cleavage and deprotect component to a purification component.

175. A system comprising a plurality of networked nucleic acid synthesizers, one or more of said networked nucleic acid synthesizer comprising the nucleic acid synthesizer of Claim 145.

176. The system of Claim 175, further comprising a dispensing component that dispenses reagents to said plurality of networked nucleic acid synthesizers.

177. The system of Claim 176, wherein said dispensing component comprises a plurality of reagent supply tanks fluidicly connected to said plurality of networked nucleic acid synthesizer, said tanks containing nucleic acid synthesis reagents, wherein at least one of said reagent supply tanks comprises at least 200 liters of acetonitrile, at least 200 liters of deblocking solution, at least 2 liters of amidite; at least 20 liters of tetrazole, at least 20 liters of capping solution, or at least 20 liters of oxidizers

178. The system of Claim 177, wherein said reagent supply tanks are contained in a first room and said plurality of nucleic acid synthesizers are contained in a second room.

179. The system of Claim 176, wherein said dispensing component comprises:

a. a plurality of valves for controlling dispensing of a plurality of reagent solutions; and

b. a plurality of dispense lines wherein each of the plurality of the dispense lines is coupled to a corresponding one of the plurality of valves for delivering one of the plurality of reagent solutions to a selected synthesis column.

180. A nucleic acid synthesizer comprising a plurality of synthesis columns and a mixing component that mixes reagents in said plurality of synthesis columns.

181. The nucleic acid synthesizer of Claim 180, wherein said mixer is selected from the group consisting of an ultrasonic mixer, a magnetic mixer, a fluid oscillator, and a vibrational mixer.

182. The nucleic acid synthesizer of Claim 180, further comprising an energy input component that imparts energy to said plurality of synthesis columns to increase nucleic acid synthesis reaction rate in said plurality of synthesis columns.

183. An oligonucleotide synthesizer comprising a reaction chamber and a lid, wherein in an open position, said lid provides a substantially enclosed ventilated workspace.

184. A method of protecting an operator of an oligonucleotide synthesizer comprising channeling ambient air away from an operator toward an interior space of said synthesizer.

185. An apparatus comprising, in combination, an oligonucleotide synthesizer and a venting hood.

186. An apparatus configured for production of oligonucleotides, wherein said apparatus comprises a venting component configured to draw air away from a reaction chamber of said apparatus.

187. A system comprising a plurality of oligonucleotide synthesizers of Claim 183.

188. The system of Claim 5, wherein said system comprises 100 or more of said synthesizers.

189. A system comprising a plurality of apparatuses of Claim 185.

190. The system of Claim 189, wherein said system comprises 100 or more of said apparatuses.

191. A system comprising a plurality of apparatuses of Claim 186.

192. The system of Claim 191, wherein said system comprises 100 or more said apparatuses.

193. A polymer synthesizer comprising a ventilated workspace.

194. The polymer synthesizer of Claim 193, wherein said polymer synthesizer is a nucleic acid synthesizer.

195. The polymer synthesizer of Claim 193, wherein said synthesizer comprises a top enclosure, said top enclosure comprising a top plate with a ventilation opening, wherein said top enclosure is configured for attachment to a top cover of a synthesizer to form a primarily enclosed space over said top cover.

196. The polymer synthesizer of Claim 193, wherein said synthesizer comprises a base, said base comprising a primarily enclosed space and a ventilation opening.

197. A top enclosure comprising a top plate with a ventilation opening, wherein said top enclosure is configured for attachment to a top cover of a synthesizer to form a primarily enclosed space over said top cover.

198.. The top enclosure of Claim 197, wherein said top plate is configured for attachment to a ventilation tube such that air in said primarily enclosed space may be drawn through said ventilation opening into said ventilation tube.

199. The top enclosure of Claim 197, wherein said top plate further comprises an outer window, and wherein said ventilation opening is formed in said outer window.

200. The top enclosure of Claim 197, wherein said top enclosure further comprises at least four sides.

201. The top enclosure of Claim 197, wherein said top cover further comprises a ventilation slot.

202. A system comprising;

a) a ventilation tube, and

b) a lid enclosure comprising; a) a top cover with a ventilation slot, and b) a top enclosure comprising a top plate with a ventilation opening, wherein said top enclosure is attached to said top cover to form a primarily enclosed space over said top cover.

203. The system of Claim 202, further comprising a vacuum source.

204. The system of Claim 202, wherein said vacuum source comprises a centralized vacuum system.

205. A kit comprising;

a) a top enclosure comprising a top plate with a ventilation opening, wherein said top enclosure is configured for attachment to a top cover of a synthesizer to form a primarily enclosed space over said top cover, and

b) a printed material component, wherein said printed material component comprises written instruction for installing said top enclosure onto said top cover.

206. A system comprising a closed system synthesizer configured for parallel synthesis of three or more polymers.

207. The system of Claim 206, wherein said three or more polymers comprise ten or more polymers.

208. The system of Claim 207, wherein said ten or more polymers comprise 48 or more polymers.

209. The system of Claim 208, wherein said 48 or more polymers comprise 96 or more polymers.

210. The system of Claim 206, wherein said polymers comprise nucleic acid polymers.

211. The system of Claim 210, wherein said nucleic acid polymers comprise DNA.

212. The system of Claim 211, wherein said DNA comprises oligonucleotides.

213. The system of Claim 206, wherein said polymers comprise three or more distinct oligonucleotides.

214. The system of Claim 213, wherein said polymers comprise twenty or more distinct oligonucleotides.

215. The system of Claim 214, wherein said polymers comprise fifty of more distinct oligonucleotides.

216. The system of Claim 206, wherein said synthesizer is configured to produce 200 or more polymers per day.

217. The system of Claim 216, wherein said synthesizer is configured to produce 1000 or more polymers per day.

218. The system of Claim 217, wherein said synthesizer is configured to produce 2000 or more polymers per day.

219. The system of Claim 216, wherein said polymers comprise oligonucleotides.

220. The system of Claim 216, wherein said oligonucleotides have 20 or more bases.

221. The system of Claim 220, wherein said oligonucleotides are produced at a 1 μmole scale.

222. The system of Claim 220, wherein said oligonucleotides are produced at less than a 1 nmole scale.

223. A synthesizer comprising:

a. a reaction support comprising three or more reaction chambers; and
b. a plurality of reagent dispensers configured to simultaneously form closed fluidic connections with each of said reaction chambers, wherein said reagent dispensers are each configured to deliver all reagents necessary for a polymer synthesis reaction.

224. The synthesizer of Claim 223, wherein said reaction support comprises 50 or more reaction chambers.

225. The synthesizer of Claim 223, wherein said reaction support comprises 96 or more reaction chambers.

226. The synthesizer of Claim 223, wherein said reaction chambers comprise synthesis columns.

227. The synthesizer of Claim 226, wherein said synthesis columns comprise nucleic acid synthesis columns.

228. The synthesizer of Claim 223, wherein said reagent dispensers are fluidicly connected to a plurality of reagent tanks.

229. The synthesizer of Claim 228, wherein said reagent dispensers are connected to said plurality of reagent tanks through a plurality of channels.

230. The synthesizer of Claim 228, wherein said plurality of reagent tanks comprise one or more tanks selected from the group consisting of acetonitrile tanks, phosphoramidite tanks, argon gas tanks, oxidizer tanks, tetrazole tanks, and capping solution tanks.

231. The synthesizer of Claim 228, wherein said reaction support comprises a fixed reaction support.

232. The synthesizer of Claim 228, wherein said reaction support further comprises a plurality of waste channels, said waste channels in closed fluidic contact with each of said reaction chambers.

233. The synthesizer of Claim 232, further comprising a detection component, wherein said detection component detects detritylation.

234. The synthesizer of Claim 233, wherein said detection component comprises a CCD camera.

235. The synthesizer of Claim 233, wherein said detection component comprises a spectrophotometer.

236. The synthesizer of Claim 233, wherein said detection component comprises a conductivity meter.

237. The synthesizer of Claim 223, further comprising a heating component.

238. The synthesizer of Claim 223, further comprising a mixing component.

239. A synthesizer comprising:

a. a fixed reaction support comprising three or more reaction chambers; and

b. a plurality of reagent dispensers configured to simultaneously form closed fluidic connections with each of said reaction chambers.

240. A nucleic acid synthesizer configured to produce 200 or more oligonucleotide per day.

241. The synthesizer of Claim 240, configured to produce 1000 or more oligonucleotides per day.

242. The synthesizer of Claim 240, configured to produce 2000 or more oligonucleotides per day.

243. The synthesizer of Claim 240, wherein said oligonucleotides comprise 20 or more bases.

244. The synthesizer of Claim 243, wherein said oligonucleotides comprise 40 or more bases.

245. The synthesizer of Claim 240, wherein said oligonucleotides are produced at a 1 μmole or greater scale.

246. The synthesizer of Claim 240, wherein said oligonucleotides are produced at a 1 nmole or smaller scale.

247. The synthesizer of Claim 240, wherein each of said 200 or more oligonucleotides comprises a different sequence.

248. A system comprising a processor, wherein said processor is configured to operate a closed system nucleic acid synthesizer for parallel synthesis of three or more nucleic acid molecules.

249. A system comprising a processor, wherein said processor is configured to operate a nucleic synthesizer and a cleavage and deprotect component.

250. The system of Claim 249, further comprising a computer memory, said computer memory comprising nucleic acid sample order information.

251. The system of Claim 250, wherein said computer memory further comprises allele frequency information.

252. The system of Claim 250, wherein said computer memory further comprises disease association information.

253. The system of Claim 206, further comprising a fail-safe reagent delivery component for delivery of one or more reagents to said solid phase synthesizer.

254. The synthesizer of Claim 223, further comprising a fail-safe reagent delivery system for delivery of one or more reagents from said reagent dispensers.

255. The solid phase synthesizer of Claim 239 further comprising a fail-safe reagent delivery system for delivery of one or more reagents from said reagent dispensers.

256. The synthesizer of Claim 240 further comprising a fail-safe reagent delivery system for delivery of one or more reagents so said synthesizer

257. The system of Claim 249 further comprising a fail-safe reagent delivery system for delivering one or more reagents to said nucleic synthesizer.

258. A system comprising a substantially closed system synthesizer configured for parallel synthesis of three or more polymers.

259. The system of Claim 206, further comprising a heating component providing substantially uniform heat during said parallel synthesis.

260. The system of Claim 259, wherein said closed system synthesizer is under controlled pressure.

261. The system of Claim 259, wherein said heating component comprises delivery of heated reagents.

262. The synthesizer of Claim 223 further comprising a heating component providing substantially uniform heat to at least two of said three or more reaction chambers.

263. The synthesizer of Claim 262, wherein said heating component comprises delivery of heated reagents to said at least two of said three or more reaction chambers.

264. The synthesizer of Claim 239, further comprising a heating component providing substantially uniform heat to at least two of said three or more reaction chambers.

265. The synthesizer of Claim 264, wherein said heating component comprises delivery of heated reagents to said at least two of said three or more reaction chambers.

266. The synthesizer of Claim 240, further comprising a heating component providing substantially uniform heat during the production of said 200 or more oligonucleotides.

267. The system of Claim 249, further comprising a heating component for providing substantially uniform heat to reaction chambers of said nucleic synthesizer, said heating element being operated by said one or more processors.

268. The system of Claim 259, wherein said heating component provides an optimized reaction temperature for a coupling step, said optimized reaction temperature being in the range of about 20 degrees C to about 60 degrees C.

269. The synthesizer of Claim 262, wherein said heating component provides an optimized reaction temperature for a coupling step, said optimized reaction temperature being in the range of about 20 degrees C to about 60 degrees C.

270. The synthesizer of Claim 264, wherein said heating component provides an optimized reaction temperature for a coupling step, said optimized reaction temperature being in the range of about 20 degrees C to about 60 degrees C.

271. The synthesizer of Claim 266, wherein said heating component provides an optimized reaction temperature for a coupling step, said optimized reaction temperature being in the range of about 20 degrees C to about 60 degrees C.

272. The system of Claim 267, wherein said heating component provides an optimized reaction temperature for a coupling step, said optimized reaction temperature being in the range of about 20 degrees C to about 60 degrees C.

273. The system of Claim 268, further comprising a mixing component.

274. The synthesizer of Claim 269, further comprising a mixing component.

275. The synthesizer of Claim 270, further comprising a mixing component.

276. The synthesizer of Claim 271, further comprising a mixing component.

277. The system of Claim 272, further comprising a mixing component.

278. The system of Claims 273 or 277, wherein said mixing component is selected from the group consisting of an ultrasonic mixer, a magnetic mixer, a fluid oscillator, and a vibrational mixer.

279. The synthesizer of Claims 274, 275 or 276, wherein said mixing component is selected from the group consisting of an ultrasonic mixer, a magnetic mixer, a fluid oscillator, and a vibrational mixer.

280. The system of Claim 206, further comprising a reagent delivery system external to said closed system synthesizer, wherein said reagent delivery system delivers one or more reagents to said closed system solid phase synthesizer.

281. The system of Claim 280, wherein said reagent delivery system further comprises at least one large volume container comprising a reagent.

282. The system of Claim 280 wherein said reagent delivery system further comprises a plurality of large volume containers, each said large volume container comprising at least one of said reagents.

283. The system of Claim 274 in which said large volume containers store in the range of about 2 liters to about 200 liters of one or more reagents.

284. The synthesizer of Claim 223, further comprising, in combination, a reagent delivery system external to said solid phase synthesizer, in which said reagent delivery system delivers one or more reagents to said solid phase synthesizer.

285. The synthesizer of Claim 284, wherein said reagent delivery system further comprises at least one large volume container comprising a reagent.

286. The synthesizer of Claim 284, wherein said reagent delivery system further comprises a plurality of large volume containers, each said large volume container comprising at least one of said reagents.

287. The synthesizer of Claim 286, wherein said large volume containers store in the range of about 2 liters to about 200 liters of one or more reagents.

288. The synthesizer of Claim 223, wherein said reagent dispensers are external to said closed system solid phase synthesizer.

289. The system of Claim 288, wherein said reagent dispensers are large volume reagent dispensers.

290. The system of Claim 289, wherein said large volume reagent dispensers store in the range of about 2 liters to about 200 liters of one or more reagents.

291. The synthesizer of Claim 240, further comprising a plurality of large volume reagent dispensers.

292. The synthesizer of Claim 291, wherein said plurality of large volume reagent dispensers are external to said synthesizer.

293. The synthesizer of Claim 292, wherein said large volume reagent dispensers store in the range of about 2 liters to about 200 liters of one or more reagents.

294. The system of Claim 248, further comprising a plurality of large volume dispensers external to said nucleic synthesizer, wherein said processor optionally regulates delivery of a reagent from said large volume dispensers.

295. The system of Claim 249, further comprising a plurality of large volume dispensers, wherein said processor optionally regulates delivery of a reagent from said large volume dispensers to said nucleic synthesizer.

296. The system of Claim 295, wherein said large volume reagent dispensers store in the range of about 2 liters to about 200 liters of one or more reagents.

297. A system comprising a substantially closed system solid phase synthesizer configured for parallel synthesis of three or more polymers; and, a fail-safe reagent delivery system for delivery of a plurality of reagents from large volume reagent delivery systems to said solid phase synthesizer.

298. The system of Claim 297, wherein said large volume reagent delivery systems are external to said solid phase synthesizer.

299. The system of Claim 298, further comprising a closed waste disposal system.

300. The system of Claim 297, further comprising one or more heaters for providing substantially uniform heat for said parallel synthesis.

301. The system of Claim 300, further comprising one or more mixers for mixing said three or more polymers.

302. A method for detecting an allele frequency of a polymorphism, comprising:

a) providing;

i) a pooled sample, wherein said pooled sample comprises target nucleic acid sequences from a plurality of individuals; and

ii) INVADER detection reagents configured to detect the presence or absence of a polymorphism; and

b) contacting said pooled sample with said INVADER detection reagents to generate a detectable signal; and

c) measuring said detectable signal, thereby determining a number of said target nucleic acid sequences that contain said polymorphism.

303. The method of Claim 302, wherein said plurality comprises at least 10 individuals.

304. The method of Claim 303, wherein said at least 10 individuals comprises at least 1000 individuals.

305. A method for detecting an allele frequency of a polymorphism, comprising:

a) providing;

i) a pooled sample, wherein said pooled sample comprises target nucleic acid sequences from a plurality of individuals; and

ii) INVADER assay detection reagents configured to generate distinct signals for each allele of a polymorphic locus in said target nucleic acid sequence;

b) contacting said pooled sample with said INVADER detection reagents to generate at least one distinct signal; and

c) measuring each of said at least one distinct signal, thereby determining a proportion of each allele of said polymorphic locus within said pooled sample.

306. The method of Claim 305, wherein said measuring comprises detection of fluorescence.

307. The method of Claim 305, wherein at least two distinct signals are generated in step b) and wherein said measuring comprises comparing said at least two distinct signals.

308. The method of Claim 306, wherein said comparing comprises applying a correction factor to a measurement of at least one distinct signal.

309. A method for detecting a rare mutation comprising;

a) providing;

i) a sample from a single subject, wherein said sample comprises at least 10,000 target nucleic acid sequences,

ii) a detection assay capable of detecting a mutation in a population of target nucleic acid sequence that is present at an allele frequency of 1:1000 or less compared to wild type alleles; and

b) assaying said sample with said detection assay under conditions such that the presence or absence of a rare mutation is detected.

310. A method for detecting a rare mutation comprising;

a) providing;

i) a sample from a single subject, wherein said sample comprises at least 10,000 target nucleic acid sequences,

ii) a detection assay capable of detecting a mutation in a population of target nucleic acid sequence that is present at an allele frequency of 1:1000 or less compared to wild type alleles; and

b) assaying said sample with said detection assay under conditions such that an allele frequency in said sample of a rare mutation is determined.

311. A panel comprising an array, said array comprising greater than about 50 different detection assays, each of said different assays being substantially similar to at least one assay shown in Figure 96.

312. The panel of Claim 311, wherein said array comprises greater than about 100 different detection assays.

313. The panel of Claim 311, wherein said array comprises greater than about 500 different detection assays.

314. The panel of Claim 311, wherein said detection assays comprise biplex detection assays.

315. The panel of Claim 311, wherein said detection assays comprise multiplex assays.

316. The panel of Claim 311, wherein said array comprises a microarray.

317. The panel of Claim 311, wherein said detection assays comprise genomic human detection assays.

318. The panel of Claim 311, wherein said array comprises a solid surface.

319. The panel of Claim 318, wherein said solid surface comprises a microtiter plate.

320. A method comprising:

a) providing:

i) a panel comprising an array, said array comprising a plurality of different assays, and

ii) a sample; and

b) exposing said sample to said panel under conditions such that at least one of said assays detects the presence of a target nucleic acid in said sample.

321. A container for conducting a plurality of detection assay reactions comprising greater than about 100 primer pairs, each pair configured to amplify a unique target sequence, said unique target sequences selected by a method comprising:

a) identifying a plurality of candidate target sequences, wherein said candidate target sequences contain a polymorphism to be detected by said detection assay; and

b) selecting said target sequences from said candidate target sequences, such that said target sequences contain only a single polymorphism detectable by said detection assay.

322. The container of Claim 321, wherein said plurality of detection assay reactions comprises greater than about 500 primer pairs.

323. The container of Claim 321, wherein said detection assay reactions comprise at least one detection assay provided in Figure 96.

324. A method of detecting a plurality of polymorphisms comprising:

a) preparing a plurality of unique target sequences from a genomic DNA sample by amplifying said target sequence using greater than about 50 different pairs of amplification primers in a single reaction container to obtain a prepared sample; and

b) exposing said prepared sample to a detection assay under conditions such that the presence or absence of a polymorphism in said unique target sequences is detected.

325. The method of Claim 14, wherein said detection assay comprises an INVADER assay.

326. The method of Claim 325, wherein said amplifying comprises no greater than 10 cycles of PCR.

327. The method of Claim 326, wherein said amplifying comprises no greater than 15 cycles of PCR.

328. The method of Claim 325, wherein said amplifying is conducted under condition such that substantially uniform amounts of each of said unique target sequences are produced.

329. The method of Claim 325, wherein said reaction container is contained on a plate.

330. A kit comprising:

a) a container comprising greater than about 50 different PCR primer pairs, and

b) a plate comprising greater than about 50 different detection assays for detecting SNPs in amplified target sequences.

331. The kit of Claim 330, wherein said detection assay is selected from the group consisting of INVADER assays, TAQMAN assays, and rolling circle assays.

332. A kit for detecting the presence of a nucleotide variation in sequence in a nucleic acid contained in a sample, which kit comprises, in packaged form, a multi-container unit having:

a. a plurality of sequence-specific detection assays functionally equivalent to assays illustrated in Figure 96; and

b. a plurality of oligonucleotide primers for each nucleic acid sequence variation being detected, said primers being selected such that an extension product synthesized from one primer, when separated from its complement, can serve as a template for the synthesis of the extension product of the other primer so as to produce amplified nucleic acid sequences containing the sequence variation.

333. A computer data storage medium comprising: a library of data for creating greater than about 100 * N assays for different single nucleotide polymorphisms, wherein N is an integer $\geq$ one.

334. The computer data storage medium of Claim 333, wherein N is an integer $\geq$ five.

335. The computer data storage medium of Claim 333, wherein said data comprises probe sequence information.

336. The computer data storage medium of Claim 335, wherein said probe sequence information comprises wild-type probe sequence information.

337. The computer data storage medium of Claim 335, wherein said probe sequence information comprises mutant probe sequence information.

338. The computer data storage medium of Claim 333, wherein said data comprises fluorescently labeled oligonucleotide data.

339. The computer data storage medium of Claim 28, wherein said fluorescently labeled oligonucleotide data comprises FRET cassette data.

340. The computer data storage medium of Claim 333, wherein said medium is selected from the group consisting of a hard drive, a floppy drive, a magnetic disk, an optical storage medium, a CD-ROM, computer memory, and a magnetic tape.

341. The computer data storage medium of Claim 333, wherein said data comprises biplex assay data.

342. The computer data storage medium of Claim 333, wherein said data comprises multiplex assay data.

343. The computer data storage medium of Claim 333, wherein said storage medium is resident on a computer.

344. The computer data storage medium of Claim 333, wherein said storage medium is resident on a plurality of computers.

345. The computer data storage medium of Claim 344, wherein said plurality of computers are communicatively linked.

346. A method for designing a detection assay comprising:

a. providing a detection assay component as shown in Figure 96; and
b. selecting an oligonucleotide sequence for a detection assay oligonucleotide using said detection assay component as a reference.

347. The method of Claim 346, wherein said selecting step comprises providing a computer memory containing a nucleic acid sequence corresponding to said detection assay component and processing said computer memory to generate said oligonucleotide sequence.

348. A method of producing a plurality of detection assays comprising:

a. providing a library of pre-validated target sequence data comprising target sequences;
b. selecting a plurality of detection assay oligonucleotide sequences complementary to said target sequences; and
c. producing a plurality of detection assays, wherein individual members of said detection assays comprise at least one oligonucleotide having said detection assay oligonucleotide sequences.

349. The method of Claim 348, wherein said detection assay comprises a TAQMAN assay.

350. The method of Claim 348, wherein said detection assay comprises a rolling circle assay.

351. The method of Claim 348, wherein said detection assay comprises a polymerase chain reaction assay.

352. The method of Claim 348, wherein said library comprises at least 100 pre-validated target sequence data of Figure 96.

353. The method of Claim 348, wherein said library comprises at least 1000 pre-validated target sequence data of Figure 96.

354. The method of Claim 348, wherein said library comprises at least 10,000 pre-validated target sequence data of Figure 96.

355. A detection assay production facility system comprising a processor configured to access the computer storage medium of Claim 333 for the production of detection assays.

356. A system comprising a library of electronic data, said data comprising: data generated during creation of (N * 1000) different SNP detection assays, where N is an interger ≥ 1.

357. The system of Claim 356, wherein N is an integer ≥ 5.

358. The system of Claim 356, wherein N is an integer ≥ 10.

359. The system of Claim 356, wherein N is an integer ≥ 20.

360. The system of Claim 356, wherein N is an integer ≥ 30.

361. The system of Claim 356, wherein N is an integer ≥ 37.

362. The system of Claim 356, wherein said data comprises pre-validated target sequence data.

363. The system of Claim 362 further comprising probe sequence data.

364. The system of Claim 356, wherein said data comprises data for greater than two different detection assay components for each different SNP assay.

365. The system of Claim 364, wherein said data comprises data for greater than three different detection assay components for each different SNP assay.

366. The system of Claim 364, wherein said data comprises data for greater than four different detection assay components for each different SNP assay.

367. The system of Claim 356, wherein said data comprises PCR primer sequence data.

368. The system of Claim 356, wherein said data comprises label data.

369. The system of Claim 356, wherein said data comprises synthetic target sequence data.

370. The system of Claim 356, wherein said data is selected from data of Figure 96 or data functionally equivalent thereto.

371. A system comprising a library of electronic data comprising at least 1000 data components representing one or more sequences shown in Figure 96.

372. The system of Claim 371 wherein said data comprises pre-validated target sequence data.

373. A method of creating a library of electronic data comprising:

a) obtaining a first electronic feed of target sequence data for a plurality of different target sequences,

b) comparing said first electronic feed of target sequence data with an electronic database of target sequence data of said different target sequences to obtain a comparison, and

c) creating a data set of pre-validated target sequences stored in said library based upon said comparison.

374. The method of Claim 373 further comprising the step of correlating probe data to said different pre-validated target sequence data.

375. The method of Claim 373 wherein said pre-validated target sequences are correlated to detection assay test result data.

376. An electronic medical record comprising, single nucleotide polymorphism data of a subject correlated to electronic medical history data of said subject.

377. The electronic medical record of Claim 376, wherein said electronic medical history data of said subject comprises prescription data.

378. The electronic medical record of Claim 377, wherein said prescription data comprising drug reaction data.

379. The electronic medical record of Claim 376, wherein said single nucleotide polymorphism data comprises data derived from an in vitro diagnostic single nucleotide polymorphism detection assay.

380. The electronic medical record of Claim 376, wherein said single nucleotide polymorphism data comprises data derived from a panel comprising a plurality of single nucleotide polymorphism detection assays.

381. The electronic medical record of Claim 380, wherein said panel comprises detection assays that detect medically associated single nucleotide polymorphisms.

382. The electronic medical record of Claim 381, wherein said panel comprises a plurality of single nucleotide polymorphism detection assays that detect single nucleotide polymorphisms associated with a disease.

383. The electronic medical record of Claim 382, wherein said panel comprises a plurality of detection assays that detect polymorphisms associated with one or more medically relevant subject areas selected from the group consisting of cardiovascular disease, oncology, immunology, metabolic disorders, neurological disorders, musculoskeletal disorders, endocrinology, and genetic disease.

384. The electronic medical record of Claim 382, wherein said panel comprises a plurality of single nucleotide polymorphism detection assays associated with two or more diseases.

385. The electronic medical record of Claim 380, wherein said panel comprises a plurality of single nucleotide polymorphism detection assays that detect polymorphisms in drug metabolizing enzymes.

386. The electronic medical record of Claim 379, wherein said single nucleotide polymorphism data comprises data derived from a plurality of in vitro diagnostic single nucleotide polymorphism detection assays.

387. The electronic medical record of Claim 386, wherein said detection assays comprises two or more unique invasive cleavage assays.

388. The electronic medical record of Claim 387, wherein one or more of said two or more unique invasive cleavage assays detected at least one single nucleotide polymorphism.

389. The electronic medical record of Claim 388, wherein said at least one single nucleotide polymorphism is associated with a medical condition.

390. The electronic medical record of Claim 387, wherein said two or more unique invasive cleavage assays comprise at least 10 unique detection assays.

391. The electronic medical record of Claim 387, wherein said two or more unique invasive cleavage assays comprise at least 1000 unique detection assays.

392. The electronic medical record of Claim 387, wherein said two or more unique invasive cleavage assays comprise at least 10,000 unique detection assays.

393. The electronic medical record of Claim 387, wherein said two or more unique invasive cleavage assays comprise at least 35,000 unique detection assays.

394. The electronic medical record of Claim 376, wherein said single nucleotide polymorphism data is derived from an analyte-specific reagent assay.

395. The electronic medical record of Claim 376, wherein said single nucleotide polymorphism data is derived from at least one clinically valid detection assay.

396. The electronic medical record of Claim 376, wherein said electronic medical record is resident on an insurance company computer system.

397. The electronic medical record of Claim 376, wherein said electronic medical record is resident on a health care provider computer system.

398. The electronic medical record of Claim 397, wherein said health care provider is selected from the group consisting of a physician computer, a hospital computer, a clinic computer, and a health maintenance organization computer.

399. The electronic medical record of Claim 376, wherein said electronic medical record is resident on a government computer system.

400. The electronic medical record of Claim 376, wherein said electronic medical record is resident on a drug store computer system.

401. The electronic medical record of Claim 376, further comprising medical billing data.

402. The electronic medical record of Claim 376, further comprising insurance claim data.

403. The electronic medical record of Claim 376, further comprising scheduling data.

404. A computer system comprising the electronic medical record of Claim 376.

405. The computer system of Claim 404, wherein said computer system is configured for receiving single nucleotide polymorphism data from the Internet.

406. The computer system of Claim 404, further comprising a software application configured to receive single nucleotide polymorphism data automatically via a communications network.

407. The computer system of Claim 404, further comprising a software application for categorizing said data.

408. The computer system of Claim 404, further comprising a software application for carrying out a bioinformatics analysis routine.

409. The computer system of Claim 404, further comprising a software application for carrying out a mathematical manipulation routine.

410. A method for determining a correlation between a polymorphism and a phenotype, comprising:

a) providing:

i) samples from a plurality of subjects;

ii) medical records from said plurality of subjects, wherein said medical records contain information pertaining to a phenotype of said subjects;

iii) detection assays that detect a polymorphism

b) exposing said samples to said detection assays under conditions such that the presence or absence of at least one polymorphism is revealed; and

c) determining a correlation between said at least one polymorphism and said phenotype of said subjects.

411. The method of Claim 410, wherein said plurality of subjects comprises 1000 or more subjects.

412. The method of Claim 410, wherein said plurality of subject comprises 10,000 or more subjects.

413. The method of Claim 410, wherein said information pertaining to a phenotype comprises information pertaining to a disease.

414. The method of Claim 410, wherein said information pertaining to a phenotype comprises information pertaining to a drug interaction.

415. The method of Claim 410, wherein said medical record comprises an electronic medical record.

416. The method of Claim 410, wherein said sample comprises a blood sample.

417. The method of Claim 410, wherein said detection assay comprises a hybridization assay.

418. The method of Claim 417, wherein said hybridization assay comprises an enzyme-based hybridization assay.

419. The method of Claim 418, wherein said enzyme-based hybridization assay comprises an invasive cleavage assay.

420. The method of Claim 410, wherein said polymorphism comprises a single nucleotide polymorphism.

421. An electronic library comprising a plurality of electronic medical records for different subjects, each of said electronic medical records comprising, single nucleotide polymorphism data of said subject correlated to electronic medical history data of said subject.

422. The electronic library of Claim 421, wherein said electronic medical history data comprises prescription data.

423. The electronic library of Claim 422, wherein said prescription data comprises drug reaction data.

424. The electronic library of Claim 421, wherein said single nucleotide polymorphism data comprises data derived from one or more in vitro diagnostic single nucleotide polymorphisms detection assays.

425. The electronic library of Claim 421, wherein said single nucleotide polymorphism data comprises data derived from a panel, said panel comprising a plurality of single nucleotide polymorphisms detection assays.

426. The electronic library of Claim 425, wherein said panel comprises detection assays that detect medically associated single nucleotide polymorphisms.

427. The electronic library of Claim 426, wherein said panel comprises a plurality of single nucleotide polymorphisms detection assays that detect single nucleotide polymorphisms associated with a disease.

428. The electronic library of Claim 427, wherein said panel comprises a plurality of detection assays that detect polymorphisms associated with one or more medically relevant subject areas selected from the group consisting of cardiovascular disease, oncology, immunology, metabolic disorders, neurological disorders, musculoskeletal disorders, endocrinology, and genetic disease.

429. The electronic library of Claim 427, wherein said panel comprises a plurality of single nucleotide polymorphism detection assays associated with two or more diseases.

430. The electronic library of Claim 425, wherein said panel comprises a plurality of single nucleotide polymorphism detection assays that detect polymorphisms in drug metabolizing enzymes.

431. The electronic library of Claim 424, wherein said single nucleotide polymorphism data comprises data derived from a plurality of in vitro diagnostic single nucleotide polymorphism detection assays for each said different subject.

432. The electronic library of Claim 421, wherein said detection assays comprises two or more unique invasive cleavage assays.

433. The electronic library of Claim 432, wherein one or more of said two or more unique invasive cleavage assays detected at least one single nucleotide polymorphism.

434. The electronic library of Claim 433, wherein said at least one single nucleotide polymorphism is associated with a medical condition.

435. The electronic library of Claim 432, wherein said two or more unique invasive cleavage assays comprise at least 10 unique detection assays.

436. The electronic library of Claim 432, wherein said two or more unique invasive cleavage assays comprise at least 1000 unique detection assays.

437. The electronic library of Claim 432, wherein said two or more unique invasive cleavage assays comprise at least 10,000 unique detection assays.

438. The electronic library of Claim 432, wherein said two or more unique invasive cleavage assays comprise at least 35,000 unique detection assays.

439. The electronic library of Claim 421, wherein said single nucleotide polymorphism data for each said different subjects is derived from an analyte-specific reagent assay.

440. The electronic library of Claim 421, wherein said single nucleotide polymorphism data for each said different subjects is derived from at least one clinically valid detection assay.

441. The electronic library of Claim 421, wherein said electronic medical record for each said different subject is resident on an insurance company computer system.

442. The electronic library of Claim 421, wherein said electronic medical record for each said different subject is resident on a health care provider computer system.

443. The electronic medical record of Claim 442, wherein said health care provider is selected from the group consisting of a physician computer, a hospital computer, a clinic computer, and a health maintenance organization computer.

444. The electronic library of Claim 421, wherein said electronic medical record for each said different subject is resident on a government computer system.

445. The electronic library of Claim 421, wherein said electronic medical record for each said different subject is resident on a drug store computer system.

446. The electronic library of Claim 421, further comprising medical billing data for each said different subject.

447. The electronic library of Claim 421, further comprising insurance claim data correlated to each said different subject data.

448. The electronic library of Claim 447, further comprising social security number data correlated to each said different subject data.

449. The electronic library of Claim 421, further comprising scheduling data correlated to each said different subject.

450. A computer system comprising the electronic library of Claim 421.

451. The computer system of Claim 450, wherein said computer system is configured for securely receiving single nucleotide polymorphism data from the Internet.

452. The computer system of Claim 450, further comprising a routine to receive single nucleotide polymorphism data for each said different subject automatically via a communications network.

453. The computer system of Claim 450, further comprising a routine to receive single nucleotide polymorphism data for each said different subject from nodes of a national, regional or world-wide communications network.

454. The computer system of Claim 450, further comprising a software application for categorizing said data for said different subjects.

455. The computer system of Claim 453, further comprising a software application for carrying out a bioinformatics analysis on said data for each said different subject.

456. A method of validating a detection assay, comprising:

a) distributing one or more detection panels to a plurality of users, wherein said detection panels comprise a plurality of candidate detection assays configured for target detection;

b) collecting test result data from at least a portion of said plurality of users, wherein said test result data is generated with said detection panels; and

c) processing at least a portion of said test result data such that at least one valid detection assay is identified from said plurality of candidate detection assays

457. The method of Claim 456, further comprising step d) marketing said valid detection assay as an Analyte-Specific Reagent or an In-Vitro Diagnostic.

458. The method of Claim 456, wherein said plurality of detection assays comprise two or more unique detection assays.

459. The method of Claim 456, further comprising a distribution system, wherein said distributing is accomplished with said distribution system.

460. The method of Claim 456, wherein said distributing one or more detection panels to said plurality of users is at a reduced cost.

461. The method of Claim 456, wherein said distributing one or more detection panels to said plurality of users is at a subsidized cost.

462. The method of Claim 456, wherein said distributing one or more detection panels to said plurality of users is at no cost.

463. The method of Claim 456, wherein prior to step a), the method further comprises the step of employing one or more of said plurality of candidate detection assays to discover at least one single nucleotide polymorphism.

464. The method of Claim 463, wherein said plurality of detection assays comprise INVADER assays.

465. The method of Claim 456, wherein prior to step a), the method further comprises the step of utilizing one or more of said plurality of candidate detection assays to associate a single nucleotide polymorphism with a medical condition.

466. The method of Claim 465, wherein said plurality of detection assays comprise INVADER assay components.

467. The method of Claim 456, wherein prior to step a), the method further comprises the step of utilizing one or more of said plurality of candidate detection assays, and bioinformatic analysis, to associate a single nucleotide polymorphism with a medical condition.

468. The method of Claim 456, wherein said plurality of detection assays comprise INVADER assay components.

469. The method of Claim 468, wherein said INVADER assay components comprise an INVADER oligonucleotide, a probe, and a control target sequence.

470. The method of Claim 456, wherein said plurality of detection assays comprise TAQMAN assay components.

471. The method of Claim 470, wherein said TAQMAN assay components comprise a probe.

472. The method of Claim 456, wherein said one or more detection panels are configured for detecting a marker associated with a disease category.

473. The method of Claim 472, wherein said disease category is selected from cardiovascular disease, cancer, autoimmune disease, metabolic disorders, neurological disease, musculoskeletal disorders, and endocrine related diseases.

474. The method of Claim 456, wherein said plurality of users comprises researchers.

475. The method of Claim 456, wherein said plurality of users comprises 10 individual users.

476. The method of Claim 456, wherein said plurality of users comprises at least 200 individual users.

477. The method of Claim 456, wherein said plurality of users comprises at least 500 individual users.

478. The method of Claim 456, wherein said plurality of users comprises at least 1000 individual users.

479. The method of Claim 456, wherein said plurality of users comprises at least 10,000 individual users.

480. The method of Claim 456, wherein said plurality of detection assays comprises at least 10 unique detection assays.

481. The method of Claim 456, wherein said plurality of detection assays comprises at least 1000 unique detection assays.

482. The method of Claim 456, wherein said plurality of detection assays comprises at least 10,000 unique detection assays.

483. The method of Claim 456, wherein said plurality of detection assays comprises at least 50,000 unique detection assays.

484. The method of Claim 456, further comprising a step, after said processing step, of selling said at least one valid detection assay as an Analyte Specific Reagent (ASR).

485. The method of Claim 456, further comprising a step, after said processing step, of selling said at least one valid detection assay as an Analyte Specific Reagent (ASR) to an In-Vitro Diagnostic Manufactuer or to a non-clinical laboratory.

486. The method of Claim 456, further comprising a step, after said processing step, of selling said at least one valid detection assay as an In-Vitro Diagnostic.

487. The method of Claim 456, wherein said test result data comprises raw assay data.

488. The method of Claim 456, wherein said test result data comprises analyzed assay data.

489. The method of Claim 456, wherein said test result data comprises data resulting from testing of 10,000 separate samples.

490. The method of Claim 456, wherein said collecting comprises receiving said test result data from at least a portion of said plurality of users over a communications network.

491. The method of Claim 490, wherein said collecting further comprises storing said test result data in a database.

492. The method of Claim 491, wherein said database is part of a computer system of a service provider.

493. The method of Claim 456, wherein said collecting comprises receiving said test result data over the Internet.

494. The method of Claim 456, wherein said collecting comprises retrieving said test result data from a user's computer system over a communication network.

495. The method of Claim 494, wherein said user's computer system comprises a software application configured to receive said test result data.

496. The method of Claim 495, wherein said software application is further configured to transmit said test result data automatically via a communications network.

497. The method of Claim 456, wherein said processing comprises categorizing said test result data.

498. The method of Claim 456, where said processing comprises computer aided data analysis of said test result data.

499. The method of Claim 456, wherein said processing comprises mathematical manipulation of said test result data.

500. The method of Claim 456, wherein said processing comprises comparing said test result data to a substantially equivalent predicate assay.

501. The method of Claim 456, wherein said processing comprises mathematical manipulation of said test result data, and comparing said test result data to a substantially equivalent predicate assay.

502. The method of Claim 456, wherein at least one valid detection assay is identified as a result of being substantially equivalent to a predicate assay.

503. The method of Claim 456, wherein said processing at least a portion of said test result data generates assay validation information.

504. The method of Claim 503, further comprising step e) submitting said assay validation information to a government body charged with approving products for clinical use.

505. The method of Claim 504, wherein said government body is the Food and Drug Administration.

506. The method of Claim 504, wherein said assay validation information is part of a 510(k) application that is submitted to the Food and Drug Administration.

507. The method of Claim 505, further comprising a step of receiving approval from said Food and Drug Administration to market said at least one valid detection assay as an FDA approved In-Vitro diagnostic assay.

508. The method of Claim 507, wherein said FDA approved In-Vitro diagnostic assay is a predicate for determining substantially equivalency for other In-Vitro diagnostic assays.

509. The method of Claim 456, wherein said target comprises a DNA molecule

510. The method of Claim 456, wherein said target comprises a RNA molecule.

511. A method of developing an in-vitro diagnostic DNA or RNA analysis product comprising, running an assay through a product development funnel, in which said assay that enters said product development funnel is substantially similar to said in-vitro diagnostic DNA or RNA analysis product.

512. The method of Claim 511 in which said assay is an assay to detect a single nucleotide polymorphism.

513. The method of Claim 511 in which said product development funnel optionally comprises a discovery portion.

514. The method of Claim 511 in which said product development funnel optionally comprises a medically associated portion.

515. The method of Claim 511 in which said product development funnel optionally comprises an analyte-specific reagent portion.

516. The method of Claim 511 in which said product development funnel optionally comprises an in-vitro diagnostic portion.

517. The method of Claim 511 in which said assay comprises a chromosome specific assay.

518. The method of Claim 511 further comprising using a panel and in which said panel comprises said assay.

519. The method of Claim 518 in which said panel comprises a medically associated portion.

520. The method of Claim 519 in which said medically associated portion comprises a panel organized by disease.

521. The method of Claim 520 in which said panel organized by disease is selected from the group consisting of a cardiovascular disease panel, an oncology panel, an immunology panel, a metabolic disorders panel, a neurological disorders panel, a musculoskeletal disorders panel, an endocrinology panel, and a genetic disease panel.

522. The method of claim 511 further comprising using a panel, and in which said panel is a panel for a multiplicity of disease states.

523. A method for manufacturing and selling detection assays, comprising:

a) receiving a customer order from a customer for a desired oligonucleotide detection assay via a computer-based customer order component;

b) processing said customer order with a detection assay design component,

c) generating a finished oligonucleotide detection assay employing a detection assay production component;

d) shipping said finished oligonucleotide detection assay to said customer employing said shipping component, and

e) billing said customer with a billing component for said oligonucleotide detection assay.

524. The method of Claim 523, wherein said billing component comprises a payment receipt component for receiving payment for said oligonucleotide detection assays.

525. The method of Claim 523, wherein said computer-based customer order component comprises a client-based computer network.

526. The method of Claim 523, wherein said computer-based customer order component comprises a distributor-based computer network.

527. The method of Claim 523, wherein said computer-based customer order component comprises a web-based user interface for ordering said oligonucleotide detection assay.

528. The method of Claim 527, wherein said web-based user interface provides a detection assay locator component.

529. The method of Claim 526, wherein said detection assay locator component comprises a library of detection assay data from which said oligonucleotide detection assay can be selected.

530. The method of Claim 529, wherein said library of detection assay data comprising single nucleotide polymorphism data.

531. The method of Claim 523, wherein said detection assay production component comprises a shop floor control system.

532. The method of Claim 531, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a make-to-order routine.

533. The method of Claim 531, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a make-to-stock routine.

534. The method of Claim 531, wherein said shop floor control system is configured to direct oligonucleotide detection assay production using a fulfill-from-stock routine.

535. The method of Claim 531, wherein said shop floor control system comprises a library of detection assay data from which said plurality of detection assay can be created.

536. The method of Claim 523, wherein said detection assay production component comprises a synthesis component.

537. The method of Claim 523, wherein said detection assay production component comprises a cleave/deprotect component.

538. The method of Claim 523, wherein said detection assay production component comprises a purification component.

539. The method of Claim 523, wherein said detection assay production component comprises a dilute and fill component.

540. The method of Claim 523, wherein said detection assay production component comprises a quality control component.

541. The method of Claim 523, wherein said synthesis component comprises a plurality of oligonucleotide synthesizers.

542. The method of Claim 541, wherein said plurality of oligonucleotide synthesizers are selected from the group consisting of MOSS EXPEDITE 16-channel DNA synthesizers (PE Biosystems, Foster City, CA), OligoPilot (Amersham Pharmacia,), the 3900 and 3948 48-Channel DNA synthesizers (PE Biosystems, Foster City, CA), POLYPLEX (Genemachines), 8909 EXPEDITE, Blue Hedgehog (Metabio), MerMade (BioAutomation, Plano, Texas), Polygen (Distribio, France), and PrimerStation 960 (Intelligent Bio-Instruments, Cambridge, MA).

543. The method of Claim 523, wherein said detection assay production component comprises an inventory control component.

544. The method of Claim 523, wherein said oligonucleotide detection assay comprises an invasive cleavage assay.

545. The method of Claim 523, wherein said oligonucleotide detection assay comprises a TAQMAN assay.

546. The method of Claim 523, wherein said oligonucleotide detection assay comprises an assay selected from the group consisting of a sequencing assay, a polymerase chain reaction assay, a hybridization assay, a hybridization assay employing a probe complementary to a mutation, a microarray assay, a bead array assay, a primer extension assay, an enzyme mismatch cleavage assay, a branched hybridization assay, a rolling circle replication assay, a NASBA assay, a molecular beacon assay, a cycling probe assay, a ligase chain reaction assay, and a sandwich hybridization assay.

547. The method of Claim 523, wherein said oligonucleotide detection assay is configured to detect a sequence selected from the group consisting of a polymorphism, a

transgene, a splice junction, a mammalian sequence, a prokaryotic sequence, and a plant sequence.

548. The method of Claim 523, wherein said detection assay production component comprises an oligonucleotide detection assay design component.

549. The method of Claim 548, where said detection assay design component comprises a PCR primer creation component.

550. The method of claim 549, wherein said PCR primer creation component is configured to optimizer PCR primer concentrations.

551. The method of Claim 548, wherein said detection assay design component is configured to design a plurality of detections assays for detecting the presence of one or more polymorphisms.

552. The method of Claim 523, wherein said order entry component or said billing component comprises a differential pricing component.

553. The method of Claim 552, wherein said differential pricing component is capable of selectably pricing said detection assay based upon a predetermined category of product.

554. The method of Claim 553, wherein said predetermined category of product is selected from the group consisting of an RUO product, an ASR product, and an IVD product.

555. The method of Claim 552, wherein said differential pricing component comprises a routine that associates a predetermined price of a detection assay based upon a presentation platform selection.

556. The method of Claim 523 in which said computer based customer order entry component further comprises a consumer direct web order entry component.

557. The method of Claim 523 in which said computer-based customer order entry component provides a data feed into said detection assay production component.

558. The method of Claim 557 in which said data feed affects production of said oligonucleotide detection assays.

559. The method of Claim 557 in which said data feed comprises statistical information associated with one or more oligonucleotide detection assays.

560. The method of Claim 559 in which said statistical information is selected from the group consisting of total oligonucleotide detection assays ordered or oligonucleotide detection assay orders received; a histogram; an oligonucleotide detection assay average per consumer; an arithemetic mean; quantity of oligonucleotide detection assays, size of order of oligonucleotide detection assays; format of panel information; a mode; a median; a weighted mean; a harmonic mean; a geometric mean; a logarithmic mean; a root mean square; a root sum square, and combination thereof; a normal distribution curve, said normal distribution curve selected from the group consisting of a normal distribution curve of number of consumers, number of detection assays, quantity of oligonucleotide detection assays, quantity of oligonucleotide detection assays or a certain type; a spread; a variance; a standard deviation; a skewed distribution; a sampling; a confidence level; and, a regression analysis.

561. A method for manufacturing and selling detection assays, comprising:

a) receiving a customer order for a desired oligonucleotide detection assay via a computer-based customer order component;

b) processing said customer order with a detection assay design component, and

c) generating a finished oligonucleotide detection assay employing a detection assay production component.

Figure 1

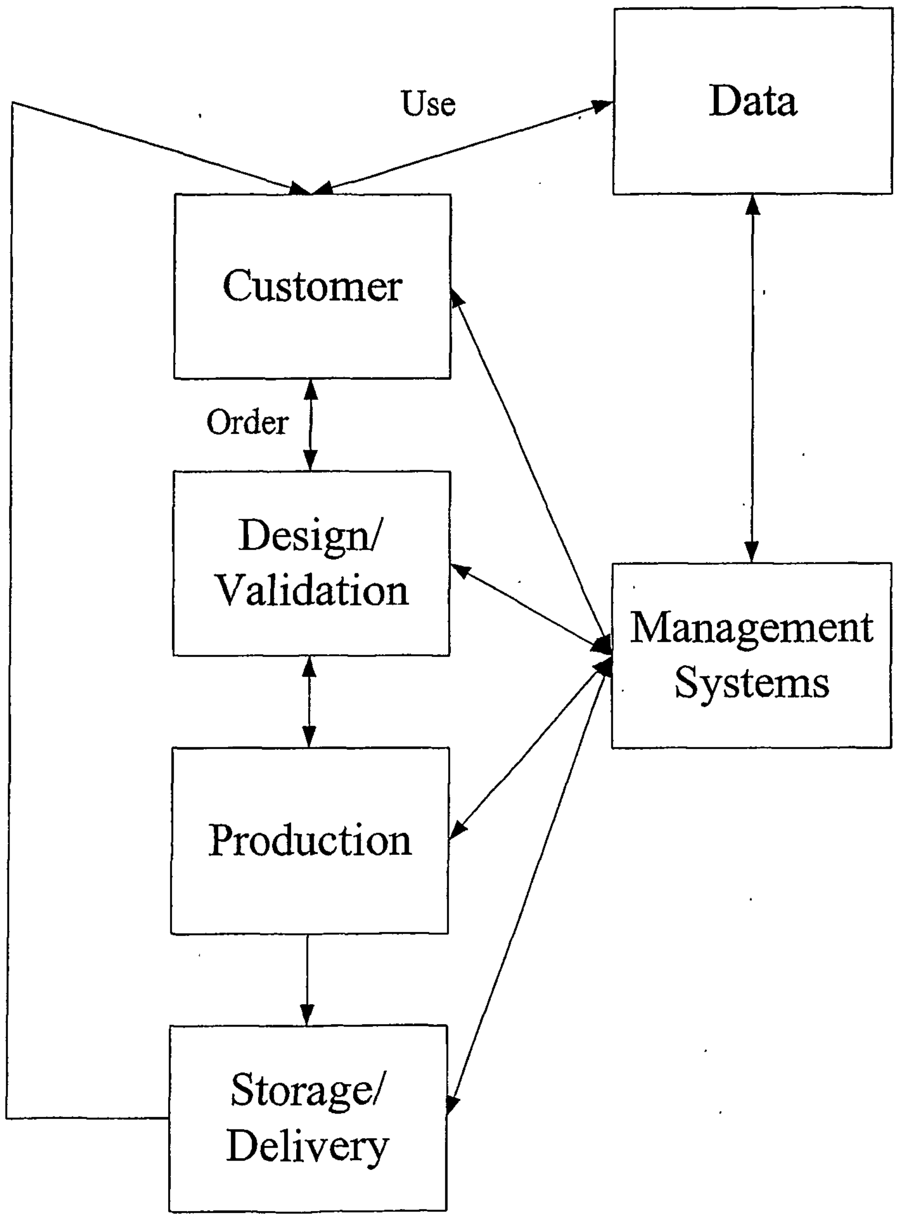

Use
Data
Customer
Order
Design/
Validation
Management
Systems
Production
Storage/
Delivery

FIGURE 2A

Invader Locator™: SNP Selection by Genome, Chromosome, Cytoband, Microsatellite Boundaries, Gene & Disease

THIRD WAVE TECHNOLOGIES INC

Home | Customer Support | Products | Order | About Us

- Invader® Locator Home
- Chromosome Map
- Gene List
- Search

Gene Name

Go

- Help

Introduction to Invader® Locator

The most extensive collection of unique assays available today to detect genetic variation in standard, custom and panel configurations. Our current portfolio of over 75 thousand SNP assays will accelerate scientific and clinical efforts related to disease gene discovery, disease susceptibility, diagnosis and treatment.

With Invader® Locator, you can choose validated SNP assays navigating at the genome, chromosome, cytoband or base pair level. You can also locate certain SNP assays by gene name or disease association, or by using a reference SNP number from these databases: NCBI's dbsnp, HGBASE, The SNP Consortium and the Japanese SNP db.

Click Chromosome Map to view genomic map.

Legal Notices | Privacy Notice | Acceptable Use Policy

FIGURE 2B

Invader Locator™: SNP Selection by Genome

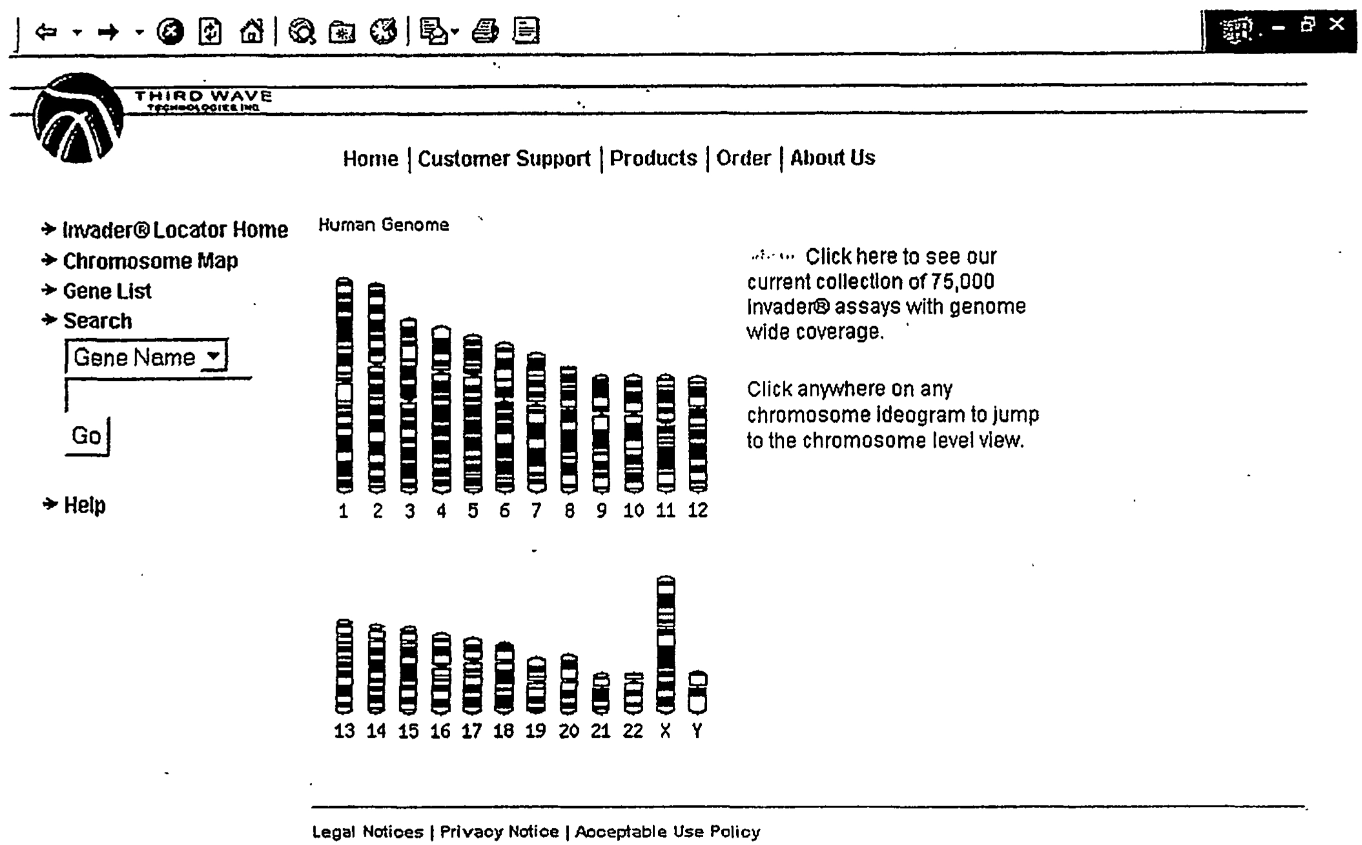


FIGURE 2C

Invader Locator™: SNP Selection by Chromosome

THIRD WAVE

Home | Customer Support | Products | Order | About Us

- Invader® Locator Home
- Chromosome Map
- Gene List
- Search

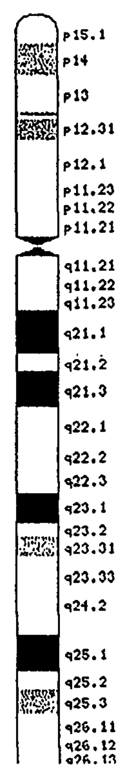
Gene Name

Go

- Help

Human Genome -> Chromosome 10

Chromosome 10



Known Genes:	483
Public SNPs:	29413
TWT SNPs:	3400
Number of Base Pairs:	144789510 bp

Click anywhere on any cytoband to jump to the cytoband level view.

FIGURE 2D

Invader Locator™: SNP Selection by Genome Cytoband, Contig, Clone & Microsatellite Boundaries

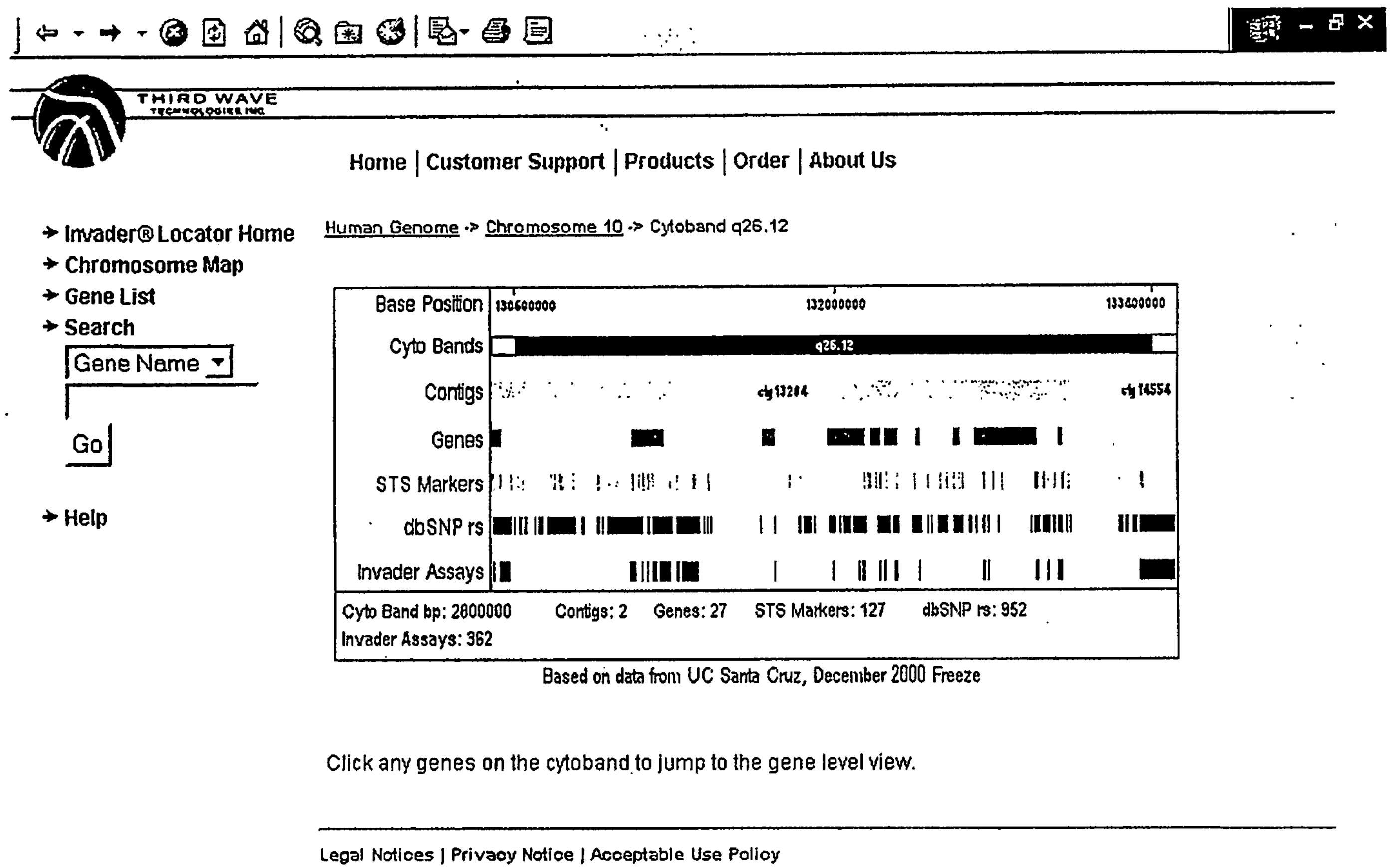


FIGURE 2E

Invader Locator™ : SNP Selection by Known & Predicted Genes

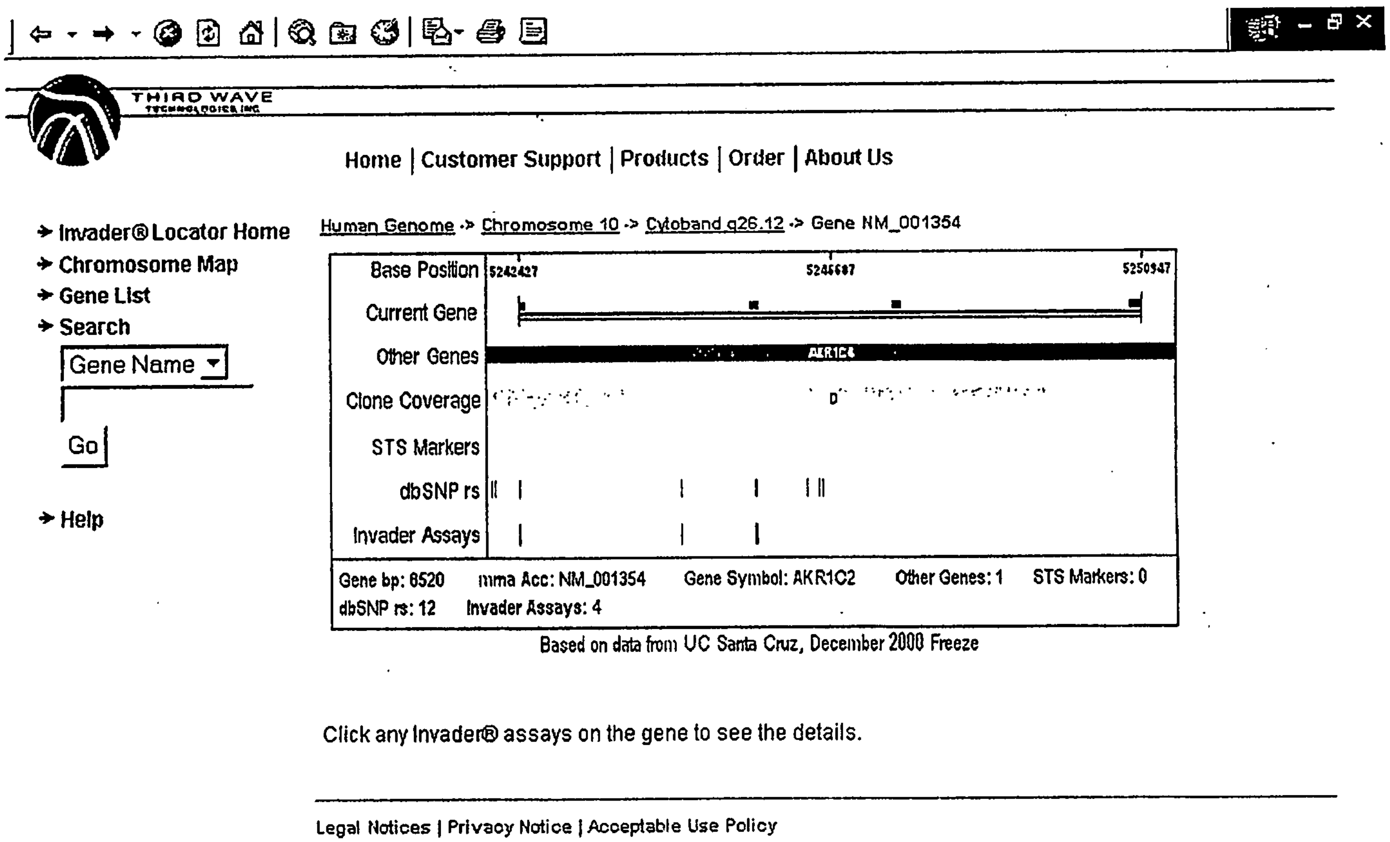


FIGURE 2F

Invader Locator™: SNP Selection by Disease Association & Phenotype

THIRD WAVE TECHNOLOGIES INC

Home | Customer Support | Products | Order | About Us

- Invader® Locator Home
- Chromosome Map
- Gene List
- Search

Gene Name

Go

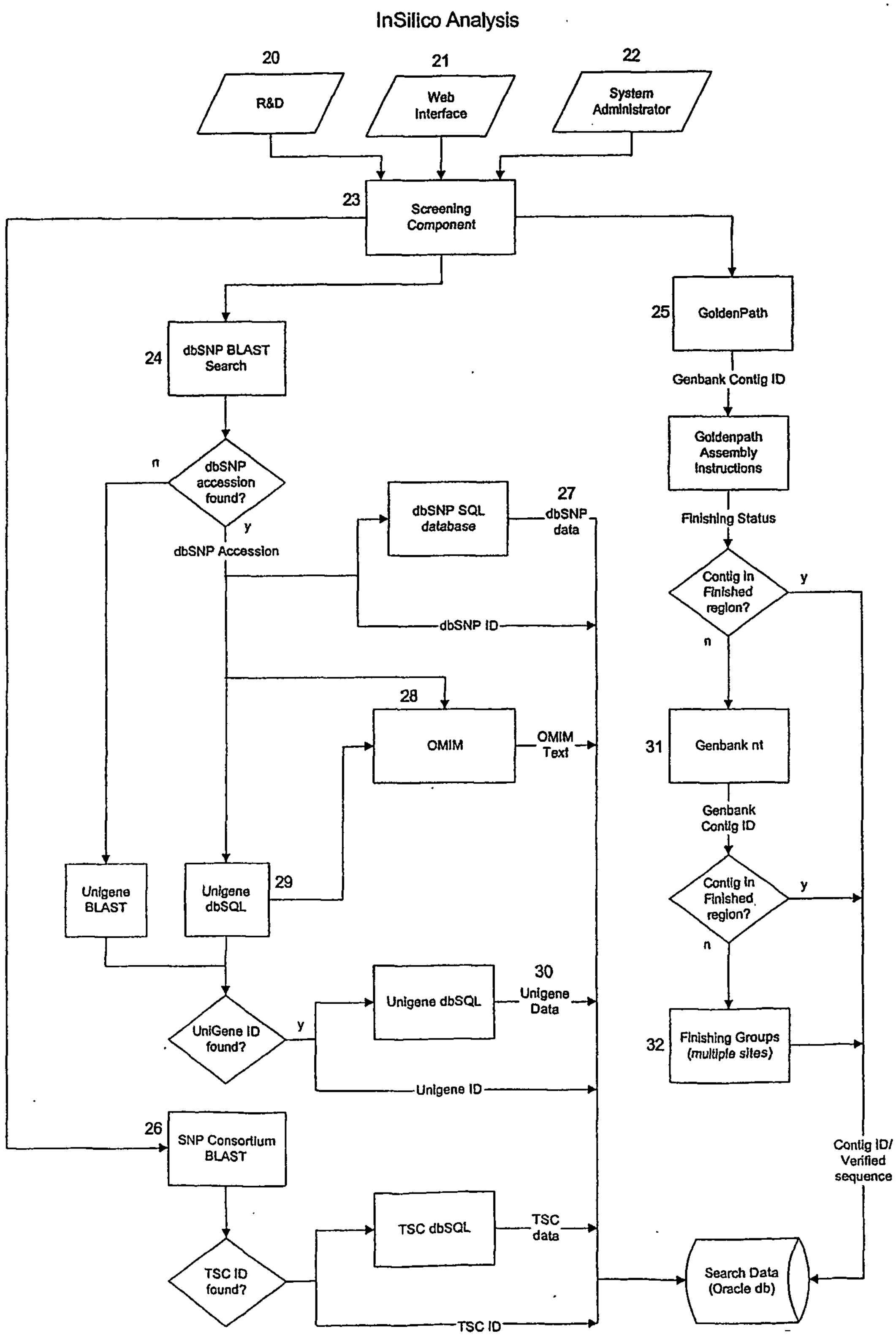
- Help

Human Genome -> Chromosome 10 -> Cytoband q26.12 -> Gene NM_001354 -> Invader® SNP 84598

Invader® SNP Number:	84598
Assay Format:	genomic
Target Sequence:	GTAATTTTGCATGATTTGAGCCATTGTT[ag] TATCTTTCGCTCTGGGAGGGAATT TTCCTACTGTTCTGAAAGGTGTCCATCACCCAAGTCACAATTGG
Genotype:	[ag]
NCBI's dbSNP rs ID:	rs8314
TSC SNP ID:	none
HGBASE SNP ID:	none
Japanese SNP ID:	none
Chromosome Location:	10p15.1
Gene Name:	aldo-keto reductase family 1, member C2
HGNC Gene Symbol:	AKR1C2
OMIM Number:	6000450
LocusLink Number:	1646
LocusLink URL:	http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?CID=1646
UniGene Link:	http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi? ORG=Hs&CID=1424
Ref mRMA:	NM_001354
Ref Protein:	NP_001345
Protein Domain:	aldo-keto reductase family 1
Ref Contig:	NT_024036

Allele Frequency | Buy this SNP

Figure 3


InSilico Analysis
20
R&D
21
Web Interface
22
System Administrator
23
Screening Component
25
GoldenPath
Genbank Contig ID
Goldenpath Assembly Instructions
Finishing Status
Contig in Finished region?
y
n
31
Genbank nt
Genbank Contig ID
Contig in Finished region?
y
n
32
Finishing Groups (multiple sites)
Contig ID/ Verified sequence
24
dbSNP BLAST Search
dbSNP accession found?
n
y
dbSNP Accession
dbSNP SQL database
27
dbSNP data
dbSNP ID
28
OMIM
OMIM Text
Unigene BLAST
Unigene dbSQL
29
UniGene ID found?
y
Unigene dbSQL
30
Unigene Data
Unigene ID
26
SNP Consortium BLAST
TSC ID found?
TSC dbSQL
TSC data
TSC ID
Search Data (Oracle db)

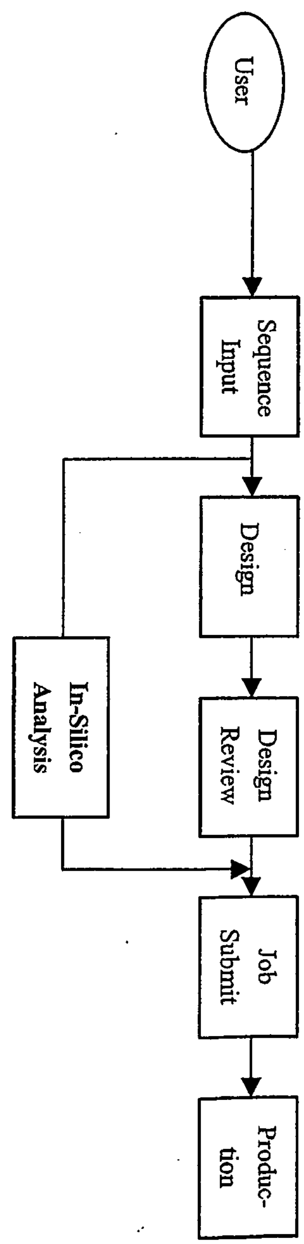

User
Sequence Input
Design
Design Review
In-Silico Analysis
Job Submit
Produc-tion

Figure 4

FIGURE 5

in silico Sequence Analysis & Database Mining

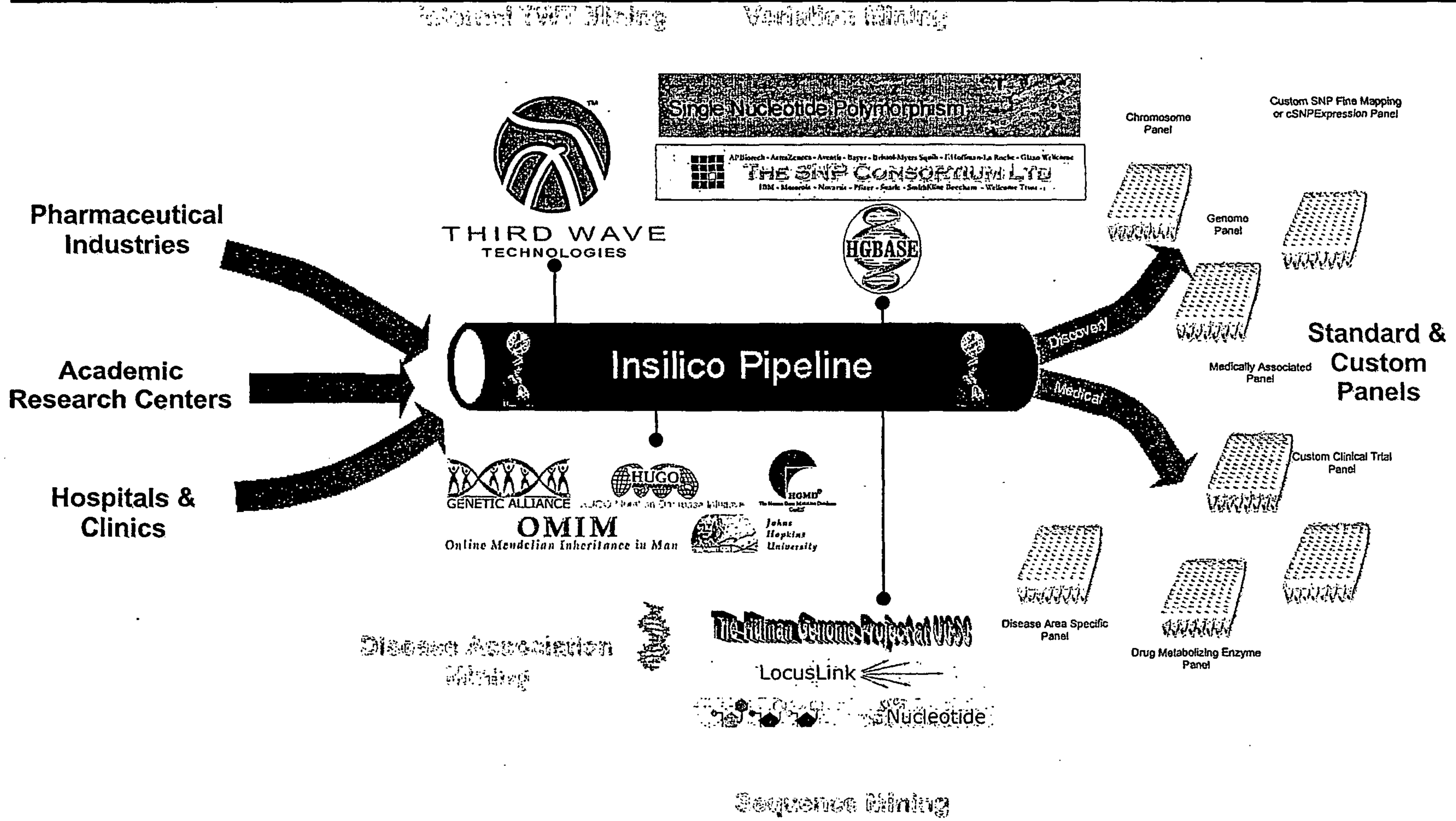
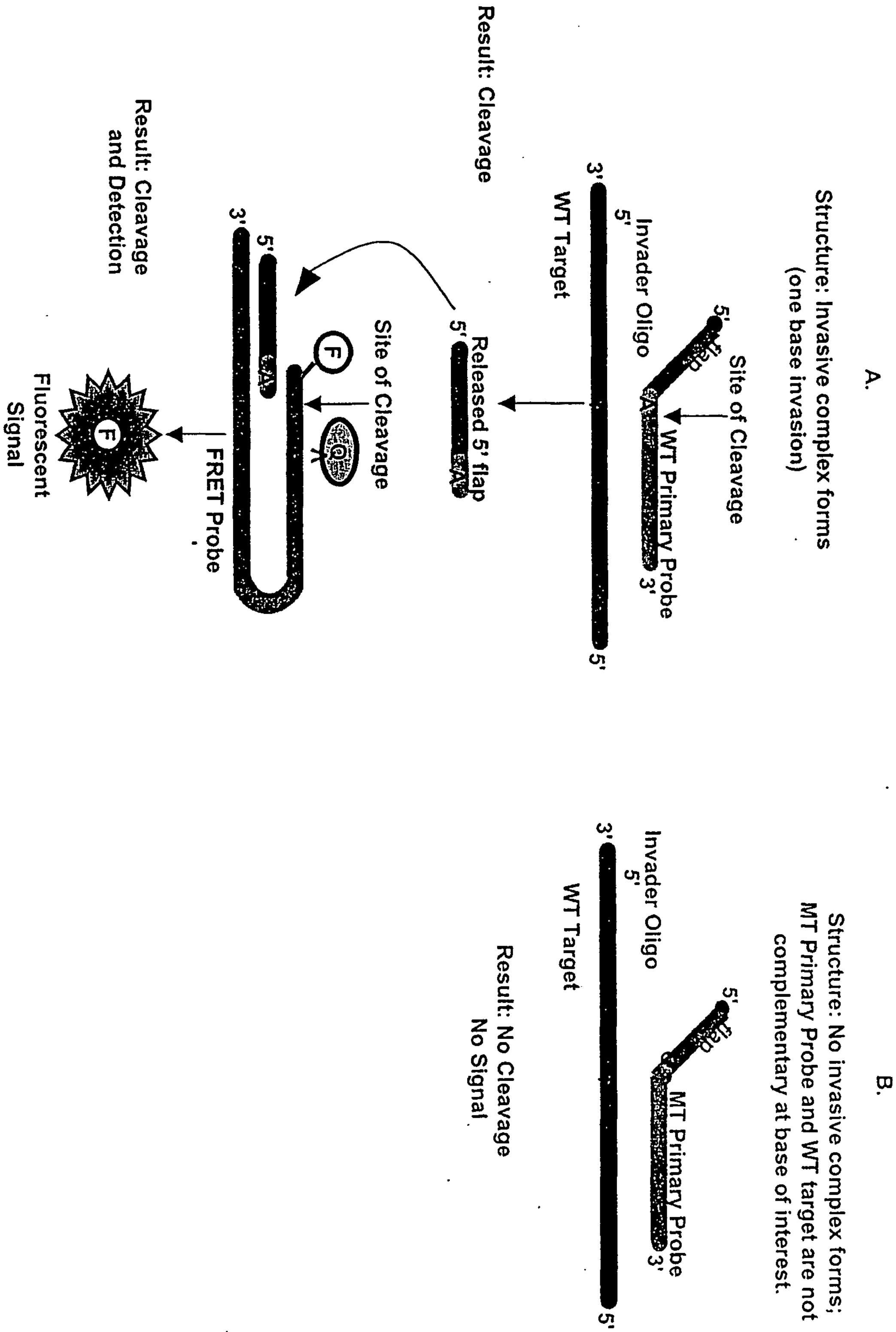


FIGURE 6
Invader Assay



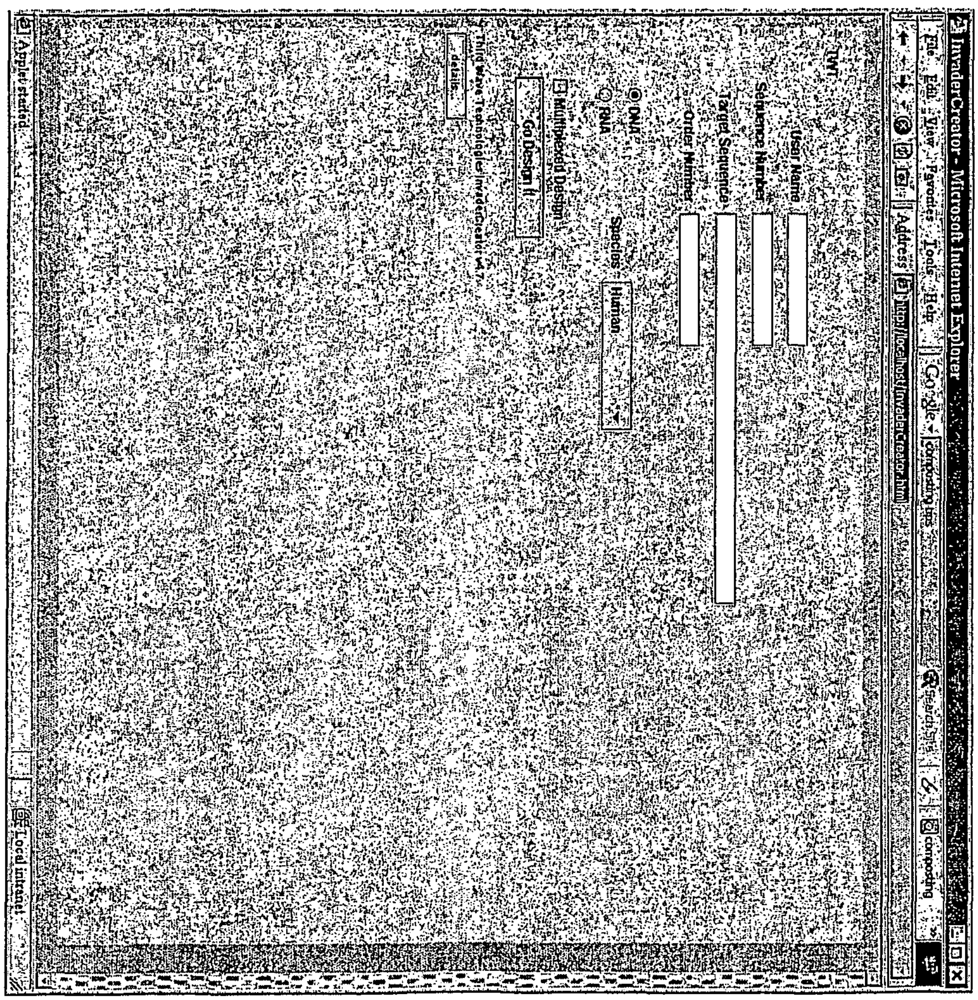

InvaderCreator - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address
User Name
Sequence Number
Target Sequence
Order Number
DNA
RNA
Species
Human
Multiplexed Design
Go Design It
details
Applet started
Local intranet

Figure 7

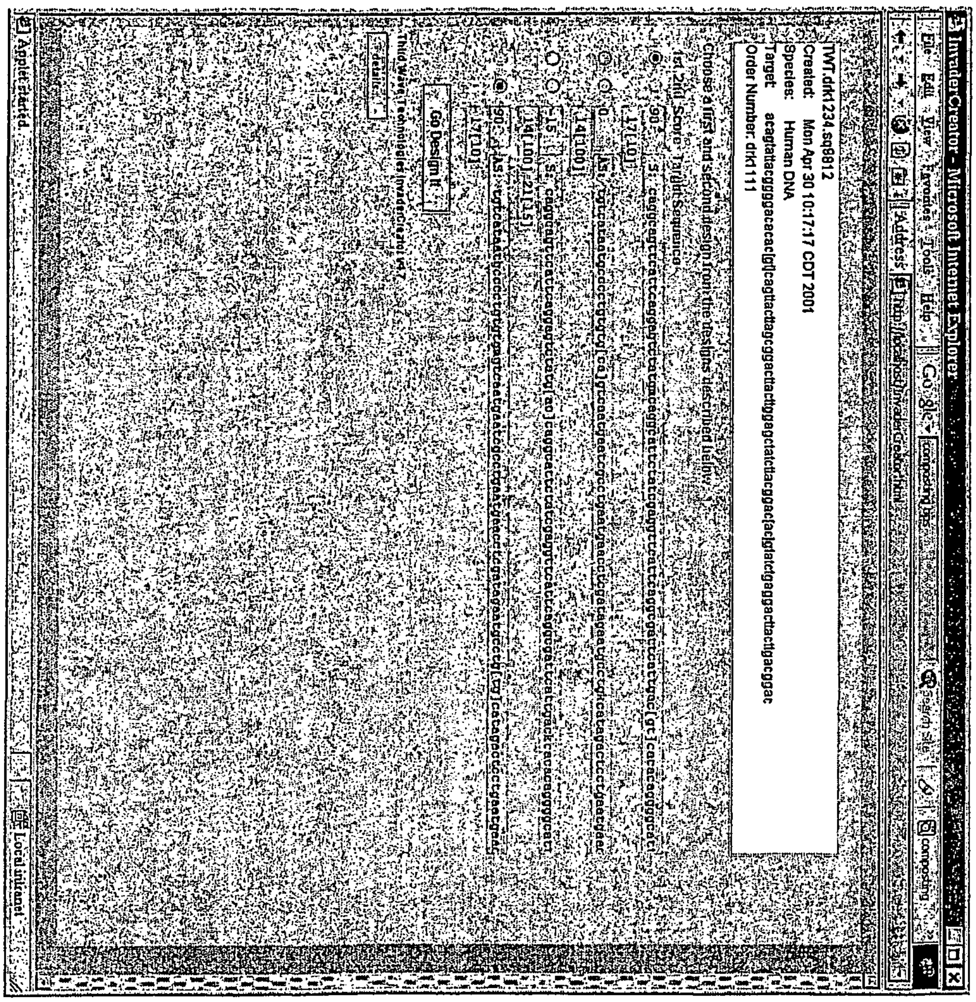

InvaderCreator - Microsoft Internet Explorer
TWT.dirk1234.sq8812
Created: Mon Apr 30 10:17:17 CDT 2001
Species: Human DNA
Order Number: dirk111
Go Design It
Applet started
Local intranet

Figure 8

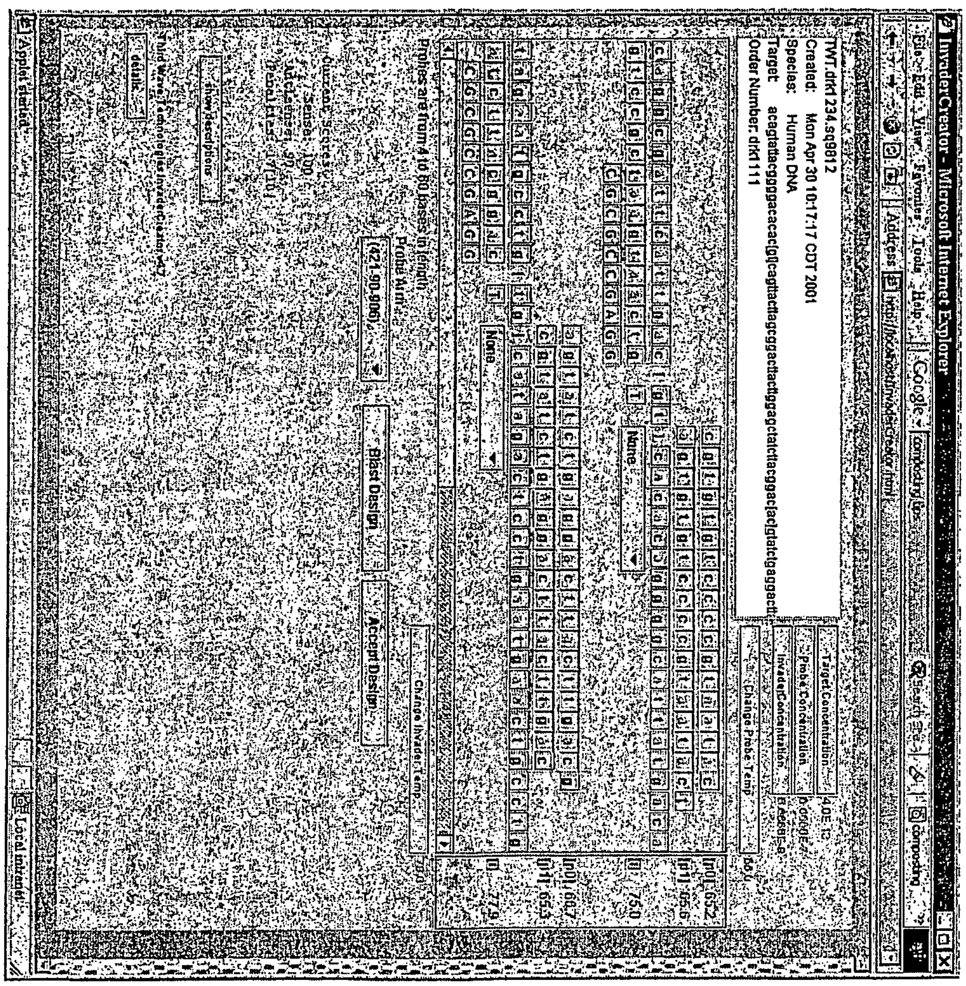

InvaderCreator - Microsoft Internet Explorer
Created: Mon Apr 30 10:17:17 CDT 2001
Species: Human DNA
Order Number: drk111
Change Probe Temp
Change Invader Temp
Blast Design
Accept Design
Applet started
Local intranet

Figure 9

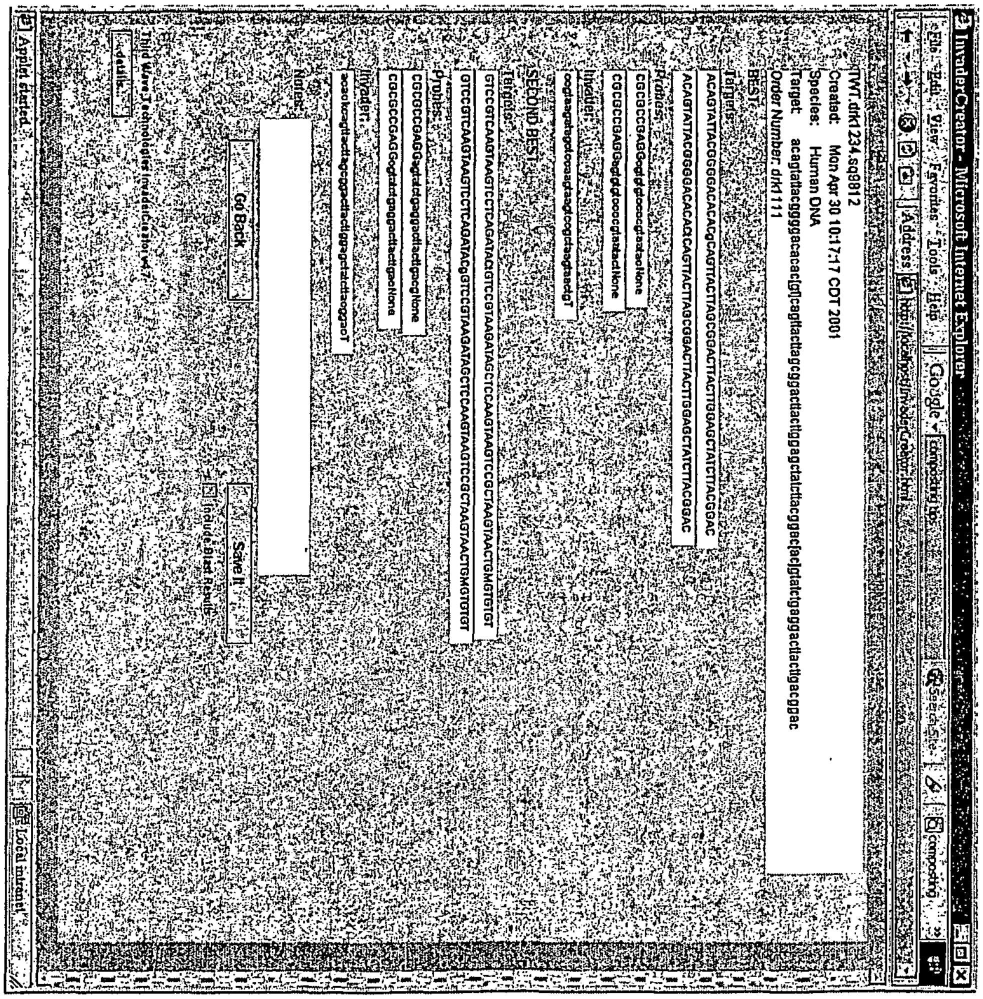


Figure 10

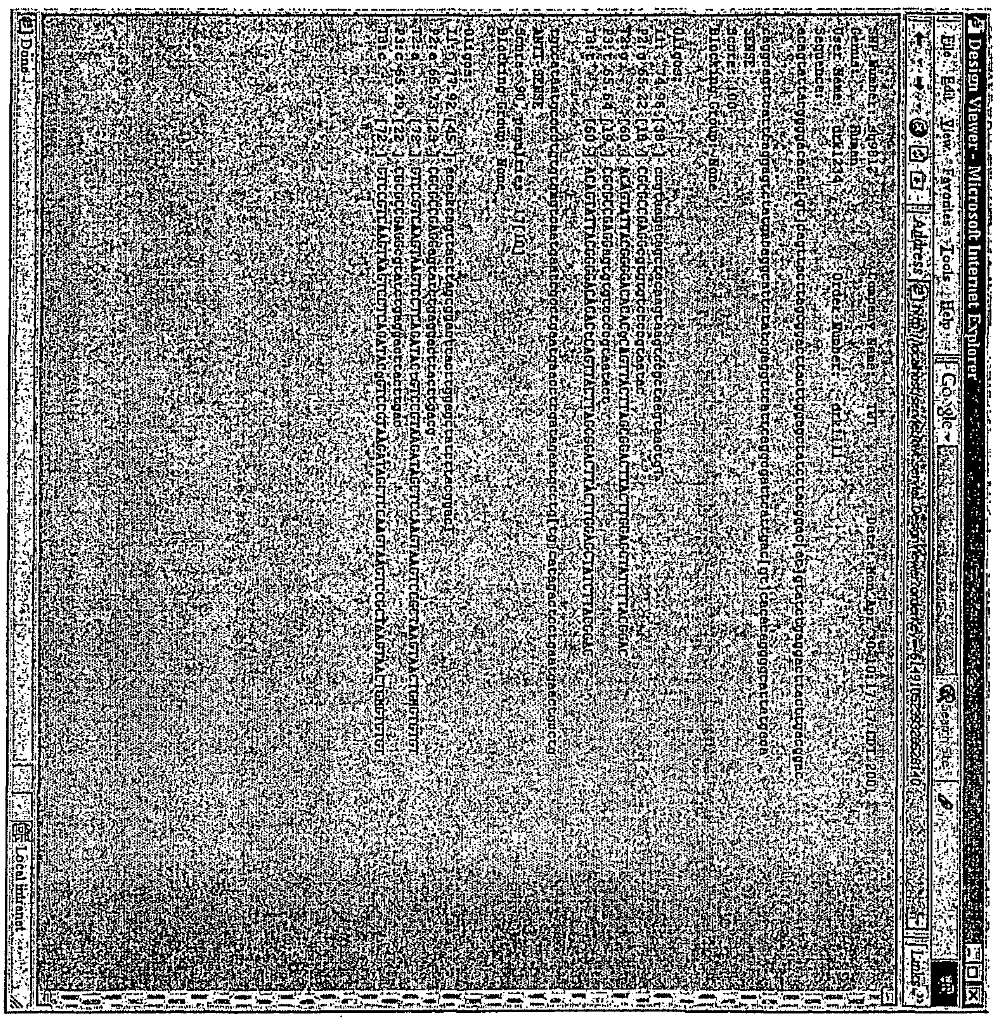


Figure 11

FIGURE 12A

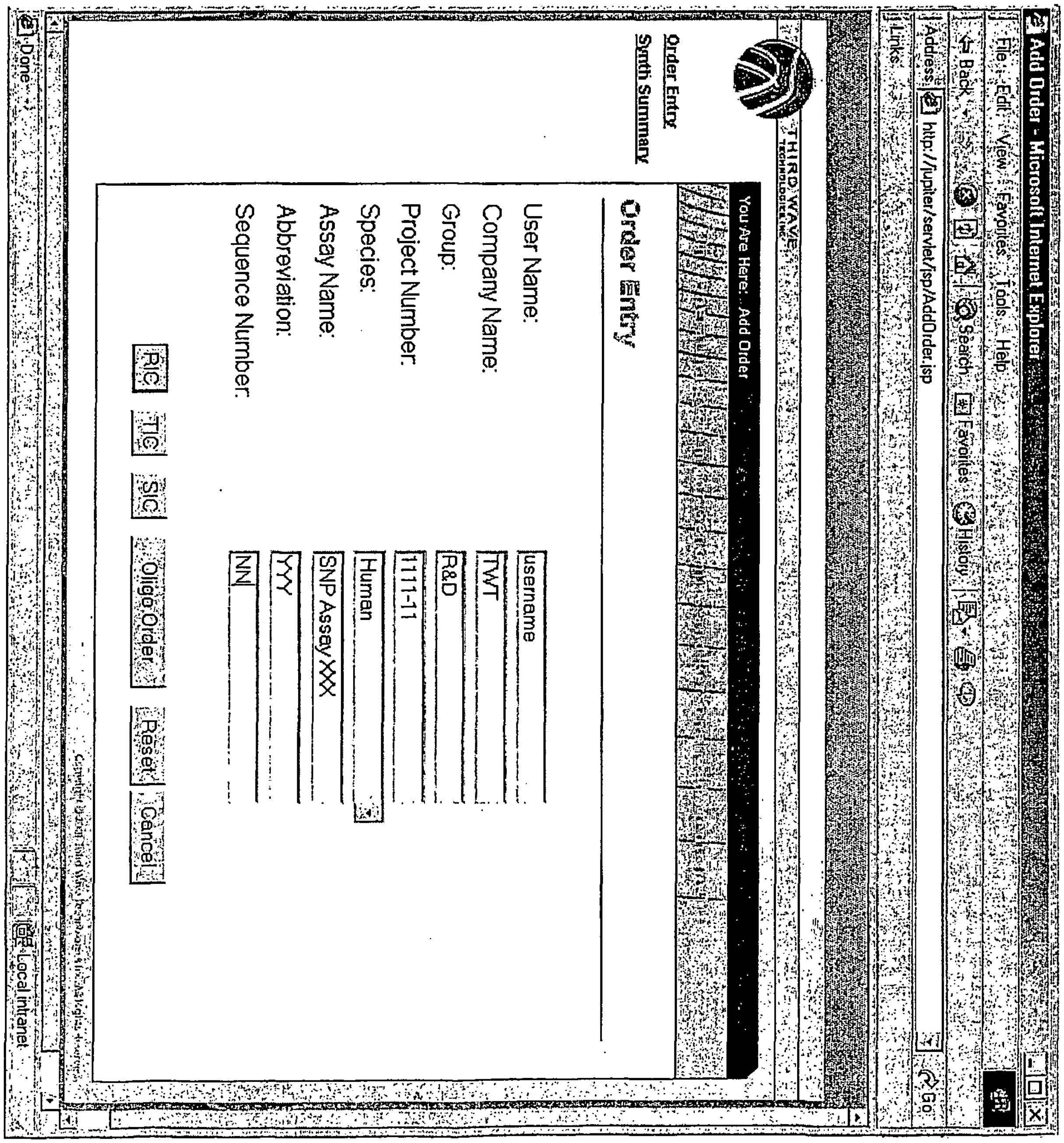

Add Order - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/AddOrder.jsp
Go
Links
THIRD WAVE TECHNOLOGIES INC.
You Are Here: Add Order
Order Entry
Synth Summary
Order Entry
User Name:
Company Name:
Group:
Project Number:
Species:
Assay Name:
Abbreviation:
Sequence Number:
username
TWT
R&D
1111-11
Human
SNP Assay XXX
YYY
NN
R|C
T|C
S|C
Oligo Order
Reset
Cancel
Done
Local intranet

FIGURE 12B

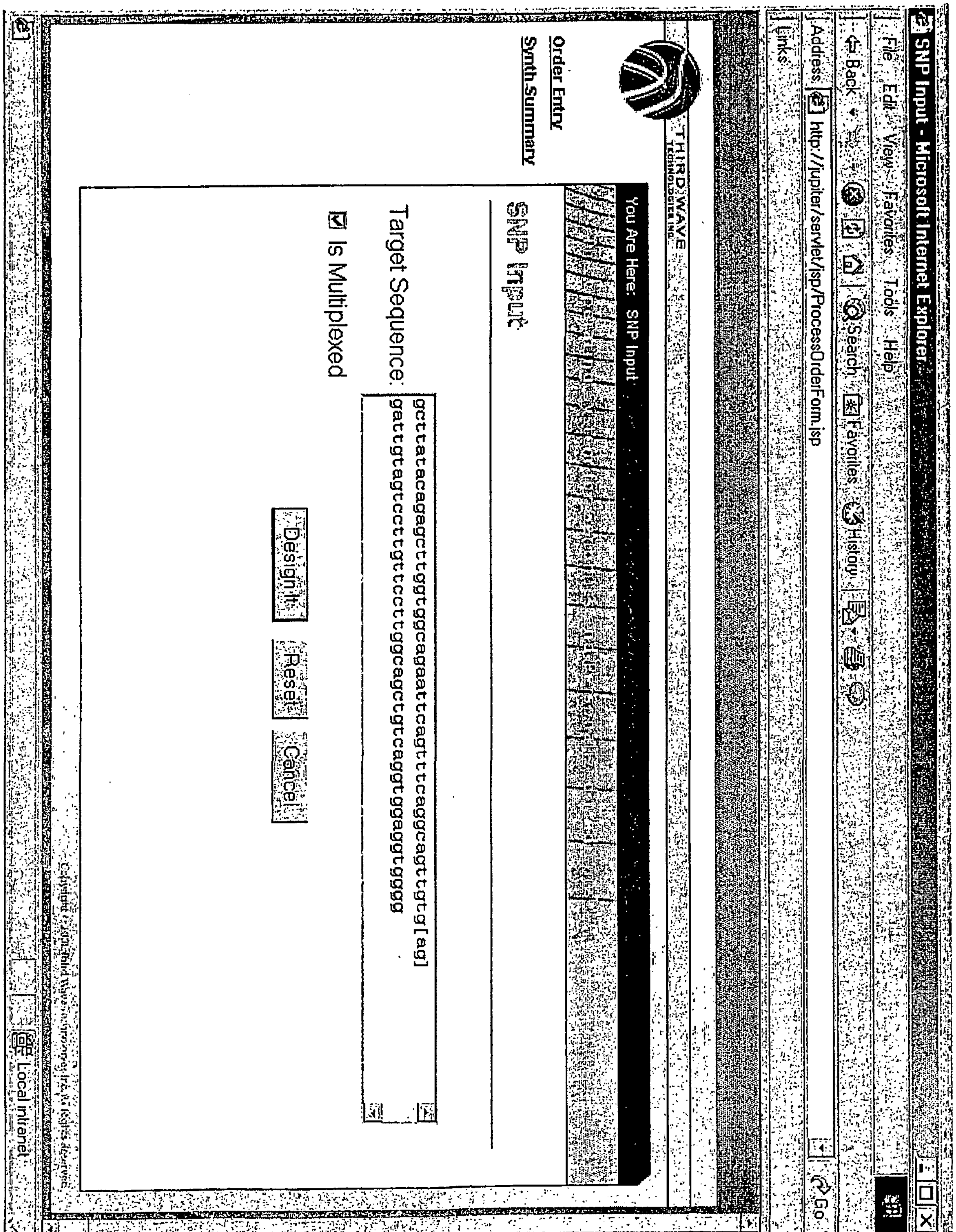

SNP Input - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessOrderForm.jsp
Go
Links
THIRD WAVE TECHNOLOGIES INC.
You Are Here: SNP Input
Order Entry
Synth Summary
SNP Input
Target Sequence:
gcttatacagagcttggtggcagaaattcagttcaggcagttgtg[ag]
gattgtagtccttgttccttggcagctgtcaggtggaggtgggg
Is Multiplexed
Design It
Reset
Cancel
Local intranet

FIGURE 12C

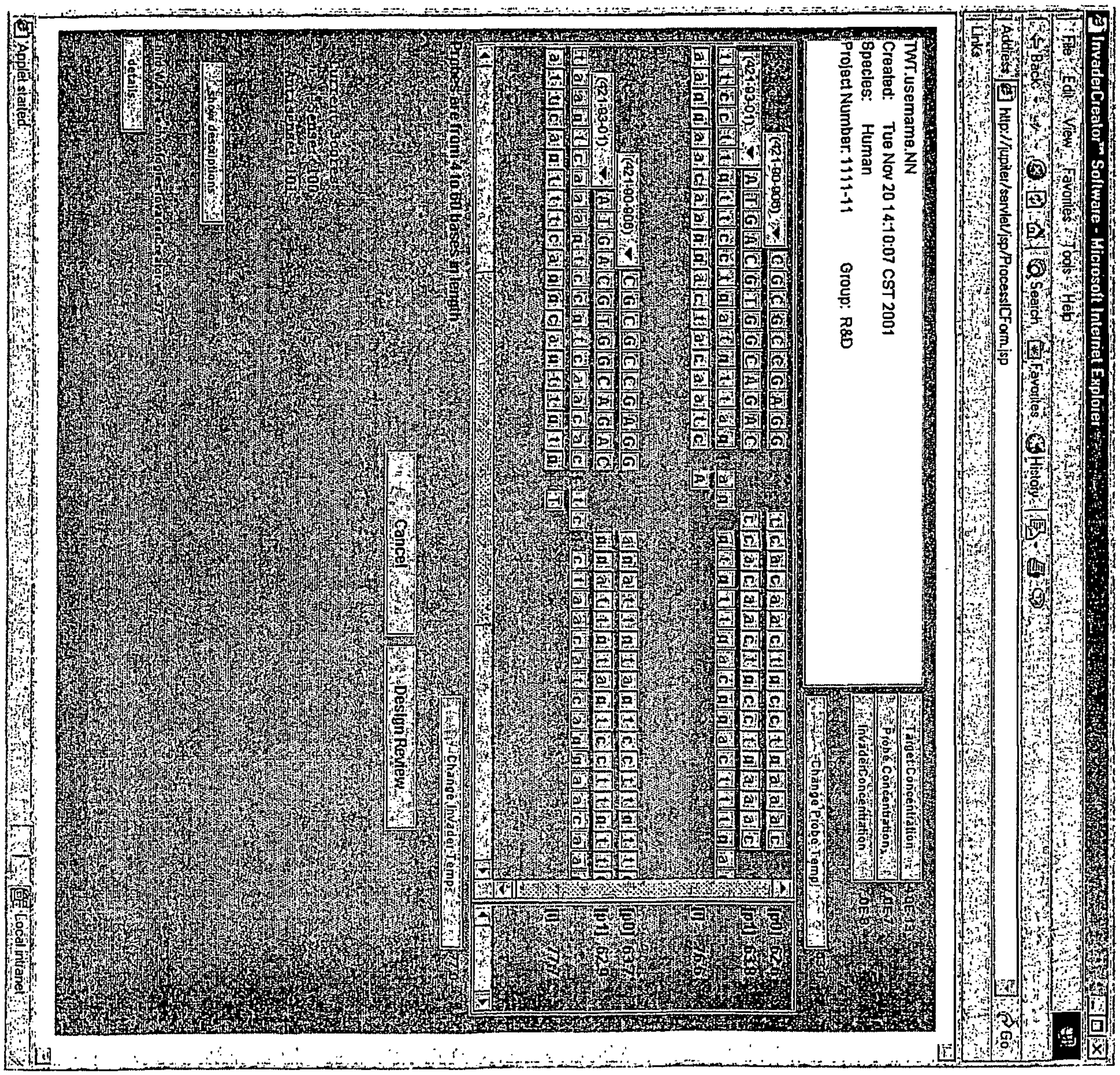
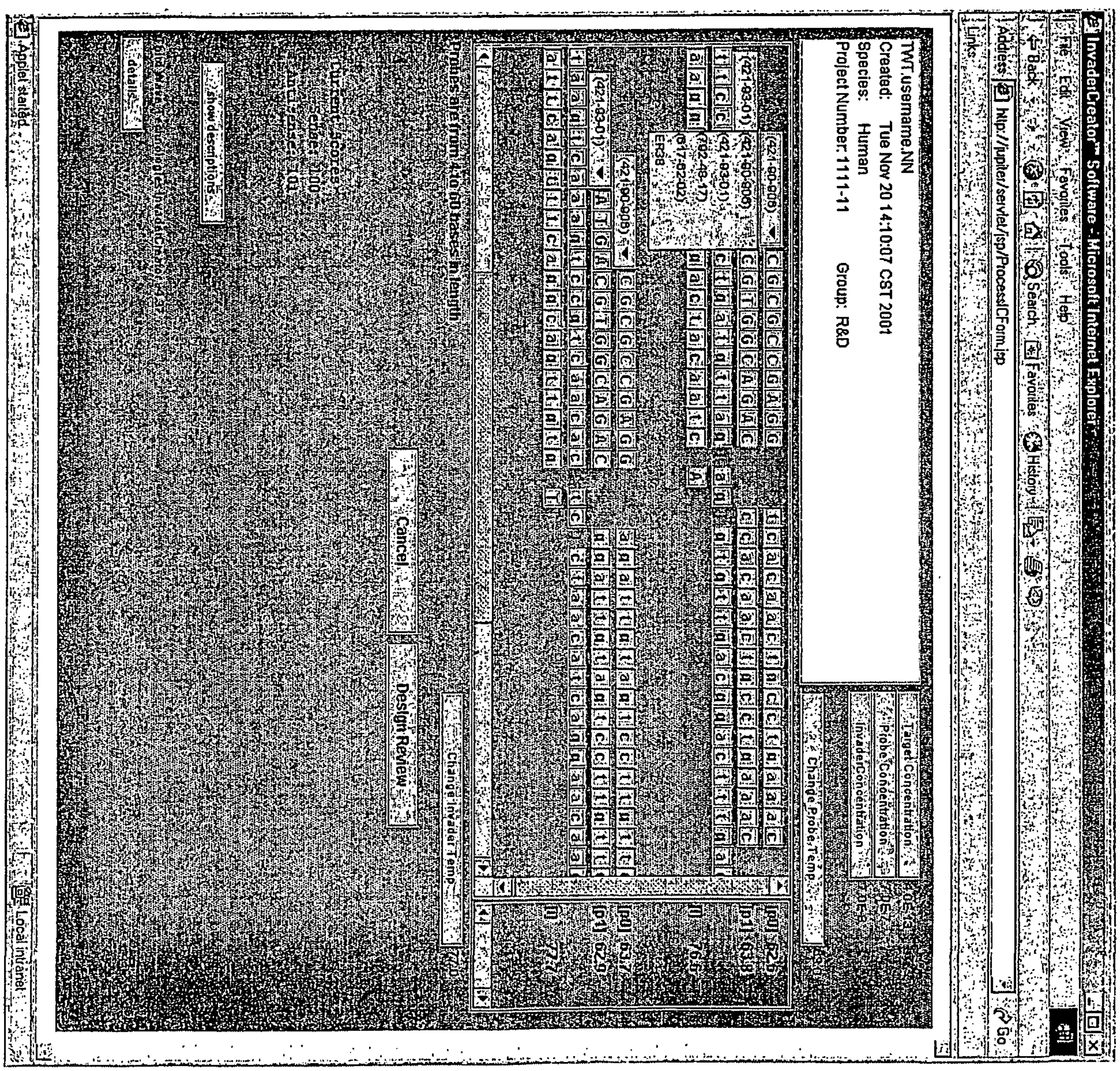

InvaderCreator™ Software - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessICForm.jsp
Links
TWT.username.NN
Created: Tue Nov 20 14:10:07 CST 2001
Species: Human
Project Number: 1111-11 Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Change Invader Temp
Cancel
Design Review
show descriptions
details
Applet started
Local intranet

FIGURE 12D


InvaderCreator™ Software - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessICForm.jsp
Go
Links
TWT.username.NN
Created: Tue Nov 20 14:10:07 CST 2001
Species: Human
Project Number: 1111-11 Group: R&D
Target Concentration
Probe Concentration
InvaderConcentration
Change Probe Temp
Change Invader Temp
Probes are from 4 to 60 bases in length
show descriptions
details
Cancel
Design Review
Applet started
Local intranet

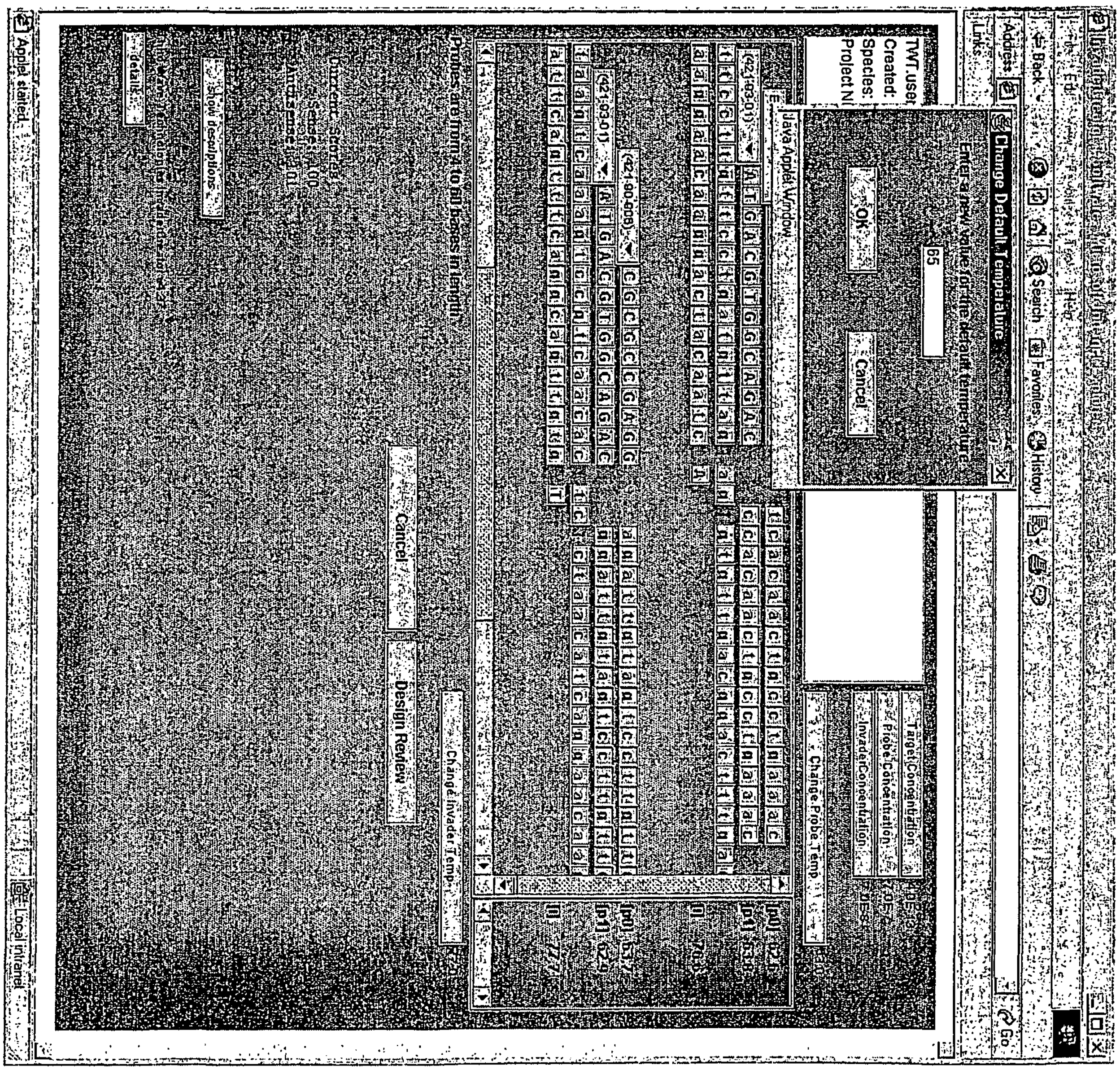

Change Default Temperature
Enter a new value for the default temperature
65
OK
Cancel
Java Applet Window
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Change Invader Temp
Design Review
Probes are from 14 to nn bases in length
show descriptions
details
Applet started
Local intranet

FIGURE 12E

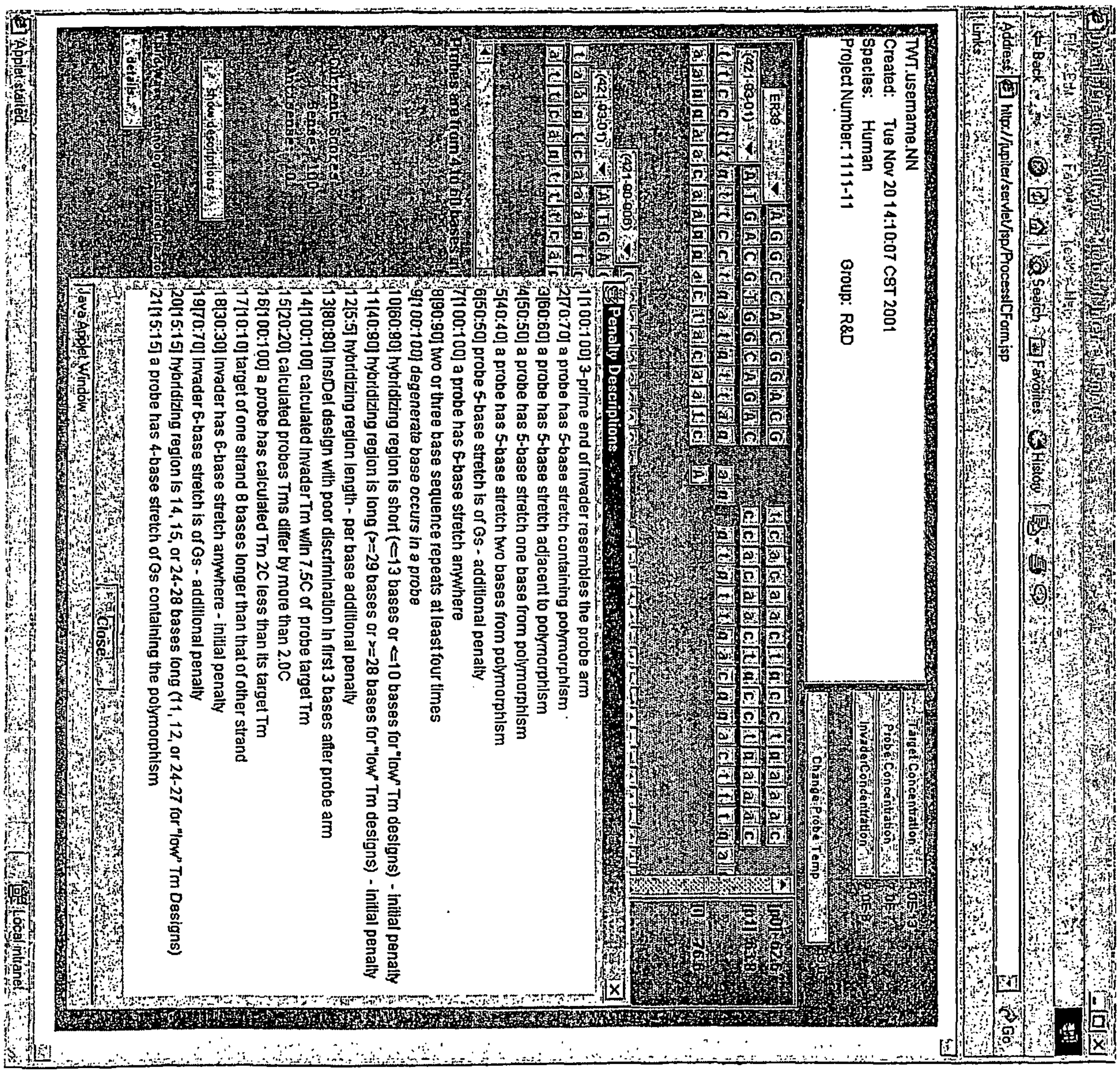

Address http://jupiter/servlet/jsp/ProcessICForm.jsp
TWT.username.NN
Created: Tue Nov 20 14:10:07 CST 2001
Species: Human
Project Number: 1111-11 Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Show descriptions
Penalty Descriptions
1[100:100] 3-prime end of invader resembles the probe arm
2[70:70] a probe has 5-base stretch containing polymorphism
3[60:60] a probe has 5-base stretch adjacent to polymorphism
4[50:50] a probe has 5-base stretch one base from polymorphism
5[40:40] a probe has 5-base stretch two bases from polymorphism
6[50:50] probe 5-base stretch is of Gs - additional penalty
7[100:100] a probe has 6-base stretch anywhere
8[90:90] two or three base sequence repeats at least four times
9[100:100] degenerate base occurs in a probe
10[60:90] hybridizing region is short (<=13 bases or <=10 bases for "low" Tm designs) - initial penalty
11[40:90] hybridizing region is long (>=29 bases or >=28 bases for "low" Tm designs) - initial penalty
12[5:5] hybridizing region length - per base additional penalty
13[80:80] Ins/Del design with poor discrimination in first 3 bases after probe arm
14[100:100] calculated invader Tm w/in 7.5C of probe target Tm
15[20:20] calculated probes Tms differ by more than 2.0C
16[100:100] a probe has calculated Tm 2C less than its target Tm
17[10:10] target of one strand 8 bases longer than that of other strand
18[30:30] Invader has 6-base stretch anywhere - initial penalty
19[70:70] Invader 6-base stretch is of Gs - additional penalty
20[15:15] hybridizing region is 14, 15, or 24-28 bases long (11, 12, or 24-27 for "low" Tm Designs)
21[15:15] a probe has 4-base stretch of Gs containing the polymorphism
Close
Java Applet Window
Applet started
Local intranet

FIGURE 12F

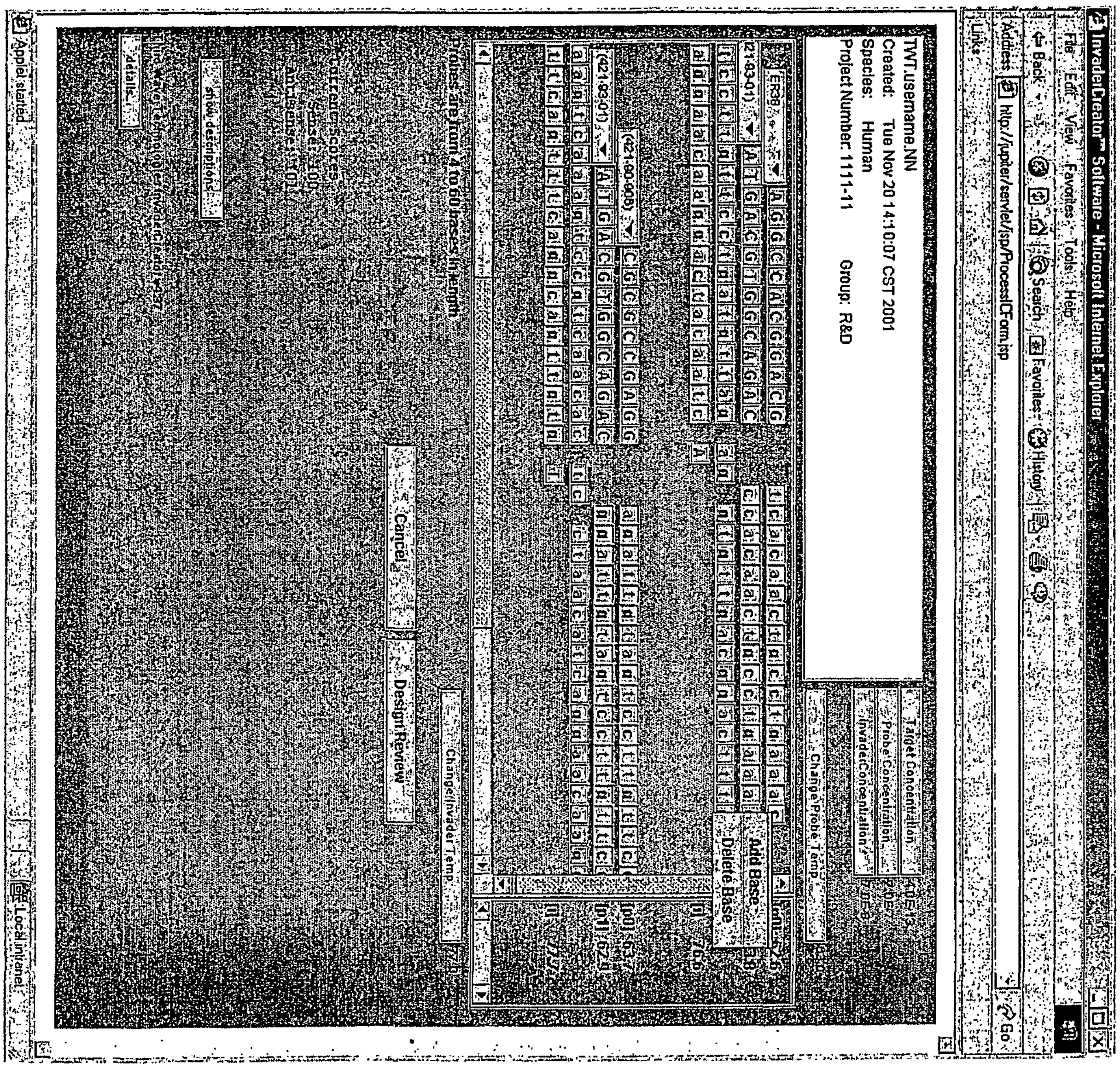

InvaderCreator™ Software - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessICForm.jsp
Links
TWT.username.NN
Created: Tue Nov 20 14:10:07 CST 2001
Species: Human
Project Number: 1111-11 Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Add Base
Delete Base
Change Invader Temp
Cancel
Design Review
show descriptions
details
Applet started
Local intranet

FIGURE 12G

FIGURE 12H

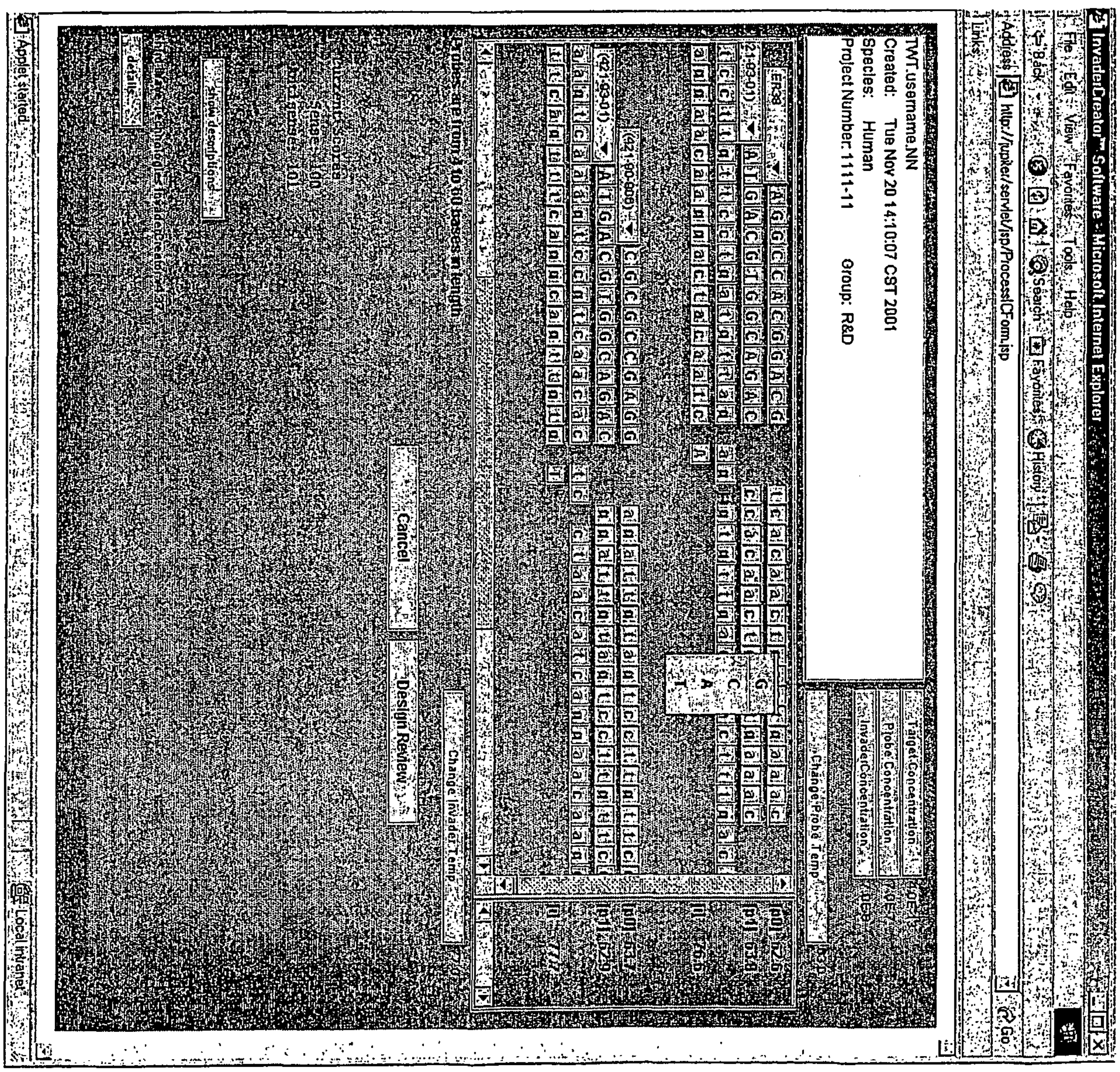

InvaderCreator™ Software - Microsoft Internet Explorer
TWT.username.NN
Created: Tue Nov 20 14:10:07 CST 2001
Species: Human
Project Number: 1111-11
Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Change Invader Temp
Cancel
Design Review

FIGURE 12I

Design Viewer - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites History

Address http://jupiter/servlet/jsp/DesignViewer.jsp?mode=SNP&orderKey=4028084799038885037 Go

Links

THIRD WAVE

Order Entry
Synth Summary

You Are Here Design Review

SNP Design Review

User Name:	username	Company Name:	TWT	Group:	R&D
Date:	11/20/01	Time:	02:10:07 CST	Design Module:	SNP
Project Number:	1111-11	Species:	Human	Sequence Number:	NN
Assay Name:	SNP Assay XXX			Abbreviation:	YYY

Sequence:
5' -> 3'

1 gcttatacagagcttggtggcagaattcagtttcaggcagttgtg[ag]gattgtagtccttgttccttggcagctgtca
78 ggtggaggtgggg

SENSE
Score: 100
Oligos:

All	Oligo Name	Type	Temp	# Bases	Sequence
☐	1111-11-1-5	Invader	76.57 °C	[34]	cctgacagctgccaaggaacaaggactacaatcA
☐	1111-11-1-1	Probe	62.62 °C	[17]	AGGCCACGGACGtcacaactgcctgaaac
☐	1111-11-1-2	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGaGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG
☐	1111-11-1-3	Probe	63.82 °C	[17]	ATGACGTGGCAGACccacaactgcctgaaac
☐	1111-11-1-4	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGgGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG

Notes:

ANTISENSE
Score: 101

Done Local intranet

Start Microsoft Excel - 1320-01 Microsoft Word - Document1 Design Viewer - Micro... 2:16 PM

FIGURE 12J

Design Viewer - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://jupiter/servlet/jsp/DesignViewer.jsp?mode=SNP&orderKey=4028084799038885037

SENSE
Score: 100
Oligos:

All	Oligo Name	Type	Temp	# Bases	Sequence
☐	1111-11-1-5	Invader	76.57 °C	[34]	cctgacagctgccaaggaacaaggactacaatcA
☐	1111-11-1-1	Probe	62.62 °C	[17]	AGGCCACGGACGtcacaactgcctgaaac
☐	1111-11-1-2	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGaGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG
☐	1111-11-1-3	Probe	63.62 °C	[17]	ATGACGTGGCAGACccacaactgcctgaaac
☐	1111-11-1-4	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGgGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG

Notes:

ANTISENSE
Score: 101
Oligos:

All	Oligo Name	Type	Temp	# Bases	Sequence
☐	1111-11-2-5	Invader	77.70 °C	[35]	gcttggtggcagaattcagtttcaggcagttgtgT
☐	1111-11-2-1	Probe	63.66 °C	[20]	CGCGCCGAGGagattgtagtccttgttcct
☐	1111-11-2-2	Synthetic Target		[60]	GCCAAGGAACAAGGACTACAATCtCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC
☐	1111-11-2-3	Probe	62.93 °C	[19]	ATGACGTGGCAGACggattgtagtccttgttcc
☐	1111-11-2-4	Synthetic Target		[60]	GCCAAGGAACAAGGACTACAATCcCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC

Notes:

Job Submit | Printable Page / Job Submit | Cancel

Done | Local intranet

Start | Microsoft Excel - 1320-01... | Microsoft Word - Document1 | Design Viewer - Micro... | 2:17 PM

FIGURE 12K

Design Viewer - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://jupiter/servlet/jsp/DesignViewer.jsp?mode=SNP&orderKey=4028084799038885037

SENSE
Score: 100
Oligos:

All	Oligo Name	Type	Temp	# Bases	Sequence
☑	1111-11-1-5	Invader	76.57 °C	[34]	cctgacagctgccaaggaacaaggactacaatcA
☑	1111-11-1-1	Probe	62.62 °C	[17]	AGGCCACGGACGtcacaactgcctgaaac
☐	1111-11-1-2	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGaGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG
☑	1111-11-1-3	Probe	63.82 °C	[17]	ATGACGTGGCAGACccacaactgcctgaaac
☐	1111-11-1-4	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGgGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG

Notes:

ANTISENSE
Score: 101
Oligos:

All	Oligo Name	Type	Temp	# Bases	Sequence
☑	1111-11-2-5	Invader	77.70 °C	[35]	gcttggtggcagaattcagtttcaggcagttgtgT
☑	1111-11-2-1	Probe	63.66 °C	[20]	CGCGCCGAGGagattgtagtccttgttcct
☑	1111-11-2-2	Synthetic Target		[60]	GCCAAGGAACAAGGACTACAATCtCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC
☑	1111-11-2-3	Probe	62.93 °C	[19]	ATGACGTGGCAGACggattgtagtccttgttcc
☑	1111-11-2-4	Synthetic Target		[60]	GCCAAGGAACAAGGACTACAATCcCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC

Notes:

Job Submit | Printable Page / Job Submit | Cancel

Done | Local intranet

Start | Microsoft Excel - 1320-01 | Microsoft Word - Document1 | Design Viewer - Micro... | 2:17 PM

FIGURE 12L

Design Viewer Printable Page - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://jupiter/servlet/jsp/DesignViewerPrint.jsp?mode=SNP&orderKey=4028084799038885037

User Name:	username	Company Name:	TWT	Group:	R&D
Date:	11/20/01	Time:	02:10:07 CST	Design Module:	SNP
Project Number:	1111-11	Species:	Human	Sequence Number:	NN
Assay Name:	SNP Assay XXX			Abbreviation:	YYY

Sequence:
5' -> 3'

1 gcttatacagagcttggtggcagaattcagtttcaggcagttgtg[ag]gattgtagtccttgttccttggcagctgtca
78 ggtggaggtgggg

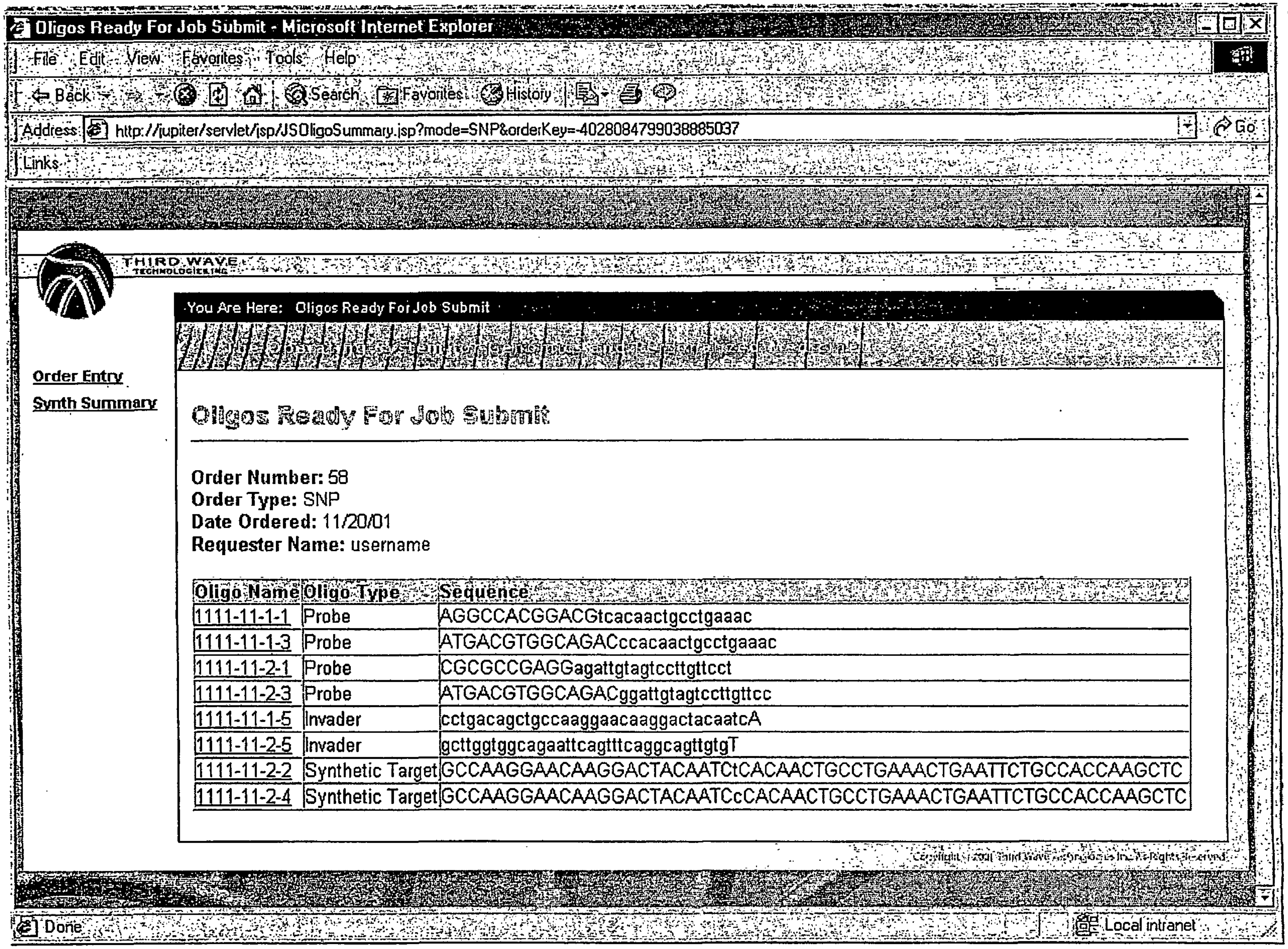
SENSE
Score: 100
Oligos:

	Oligo Name	Type	Temp	# Bases	Sequence
*	1111-11-1-5	Invader	76.57 °C	[34]	cctgacagctgccaaggaacaaggactacaatcA
*	1111-11-1-1	Probe	62.62 °C	[17]	AGGCCACGGACGtcacaactgcctgaaac
	1111-11-1-2	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGaGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG
*	1111-11-1-3	Probe	63.82 °C	[17]	ATGACGTGGCAGACccacaactgcctgaaac
	1111-11-1-4	Synthetic Target		[56]	TTCAGTTTCAGGCAGTTGTGgGATTGTAGTCCTTGTTCCTTGGCAGCTGTCAGGTG

Notes:

Done Local intranet

FIGURE 12M



Oligo Name	Oligo Type	Sequence
1111-11-1-1	Probe	AGGCCACGGACGtcacaactgcctgaaac
1111-11-1-3	Probe	ATGACGTGGCAGACccacaactgcctgaaac
1111-11-2-1	Probe	CGCGCCGAGGagattgtagtccttgttcct
1111-11-2-3	Probe	ATGACGTGGCAGACggattgtagtccttgttcc
1111-11-1-5	Invader	cctgacagctgccaaggaacaaggactacaatcA
1111-11-2-5	Invader	gcttggtggcagaattcagtttcaggcagttgtgT
1111-11-2-2	Synthetic Target	GCCAAGGAACAAGGACTACAATCtCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC
1111-11-2-4	Synthetic Target	GCCAAGGAACAAGGACTACAATCcCACAACTGCCTGAAACTGAATTCTGCCACCAAGCTC

FIGURE 12N

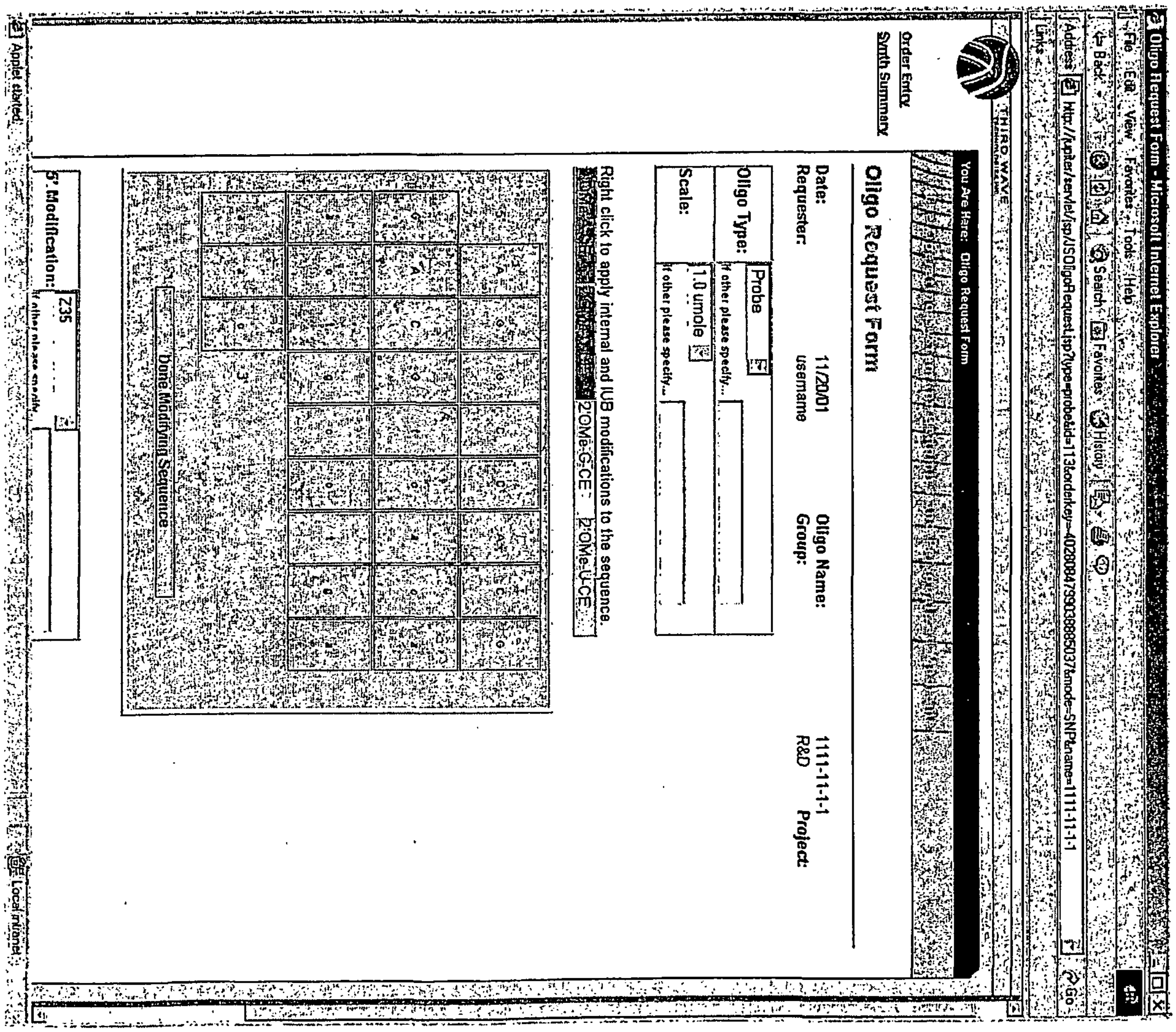

Oligo Request Form - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address
Links
THIRD WAVE
Order Entry
Synth Summary
You Are Here: Oligo Request Form
Oligo Request Form
Date: 11/2001
Requester: username
Oligo Name: 1111-11-1-1
Group: R&D
Project:
Oligo Type: Probe
If other please specify...
Scale: 1.0 umole
If other please specify...
Right click to apply internal and IUB modifications to the sequence.
2'OMe-G-CE
2'OMe-U-CE
Done Modifying Sequence
5' Modification: Z35
Applet started
Local intranet

FIGURE 120

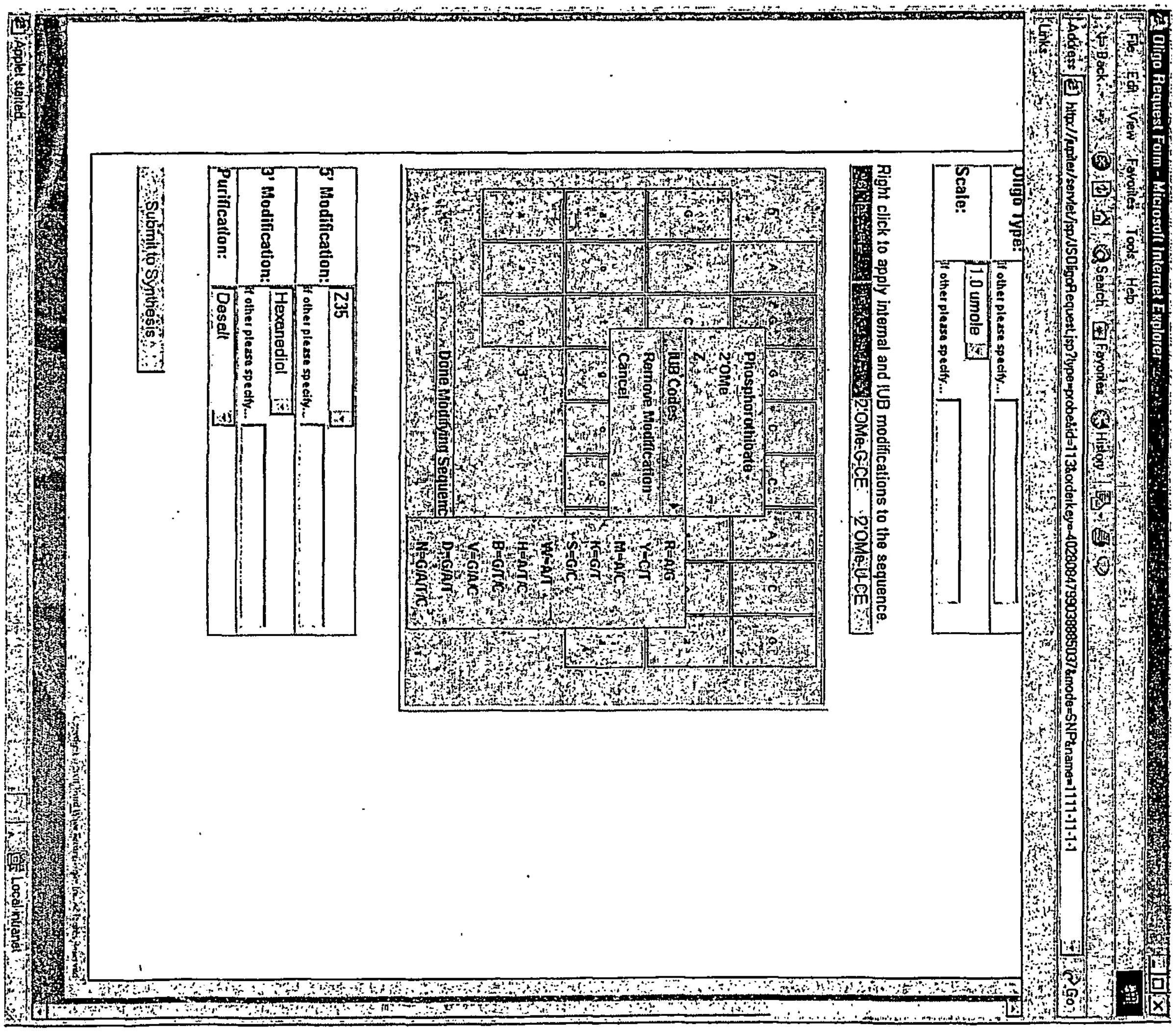

Oligo Request Form - Microsoft Internet Explorer
Oligo Type:
Scale:
1.0 umole
Right click to apply internal and IUB modifications to the sequence.
Phosphorothioate
2'OMe
IUB Codes
Remove Modification
Cancel
5' Modification:
Z35
3' Modification:
Hexanediol
Purification:
Desalt
Submit to Synthesis
Applet started
Local intranet

FIGURE 12P

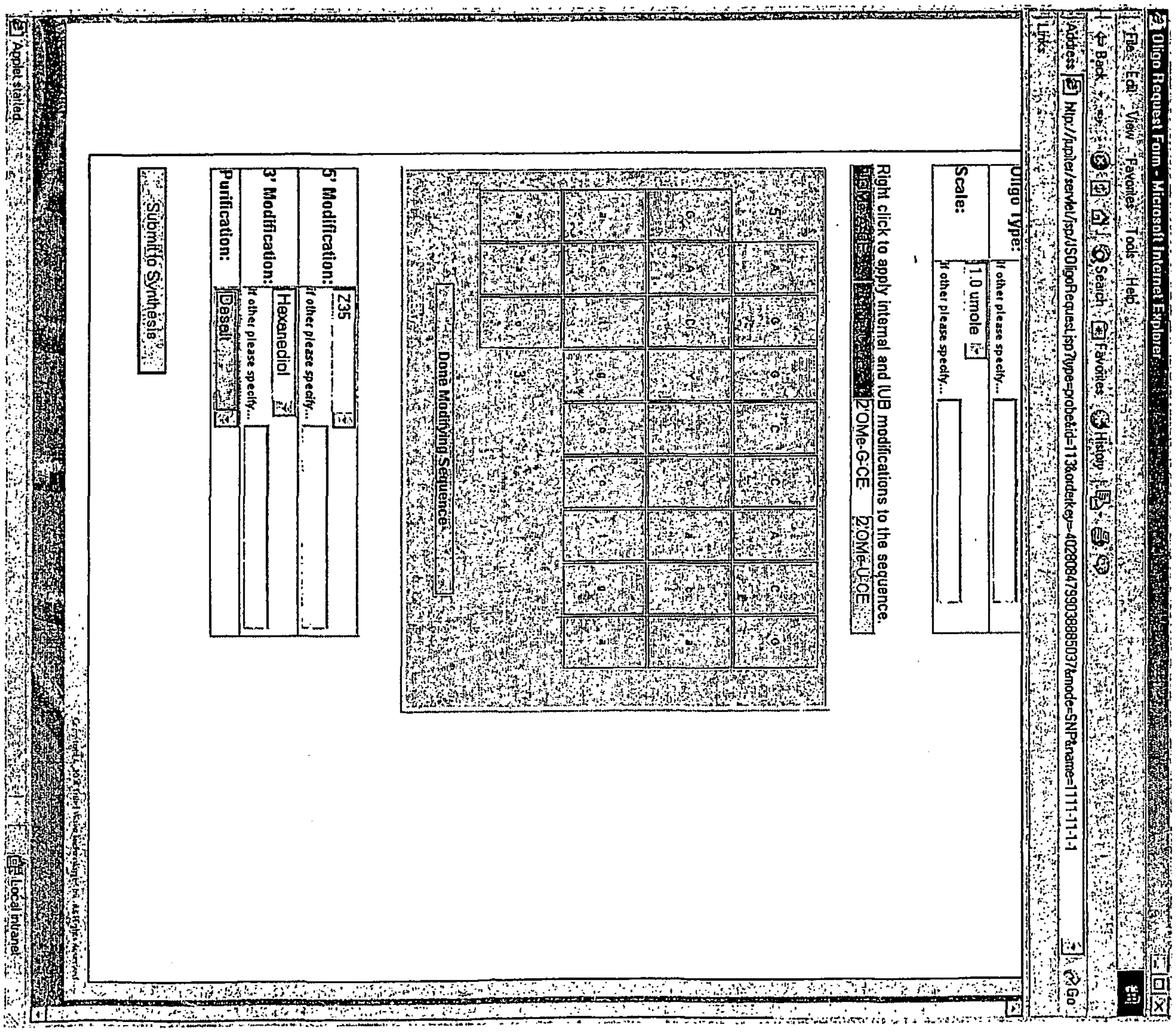

Oligo Request Form - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address
Links
Oligo Type:
Scale:
1.0 umole
If other please specify...
Right click to apply internal and IUB modifications to the sequence.
2'OMe-G-CE
2'OMe-U-CE
Done Modifying Sequence
5' Modification:
3' Modification:
Purification:
Hexanediol
Desalt
Submit to Synthesis
Applet started
Local intranet

FIGURE 12Q

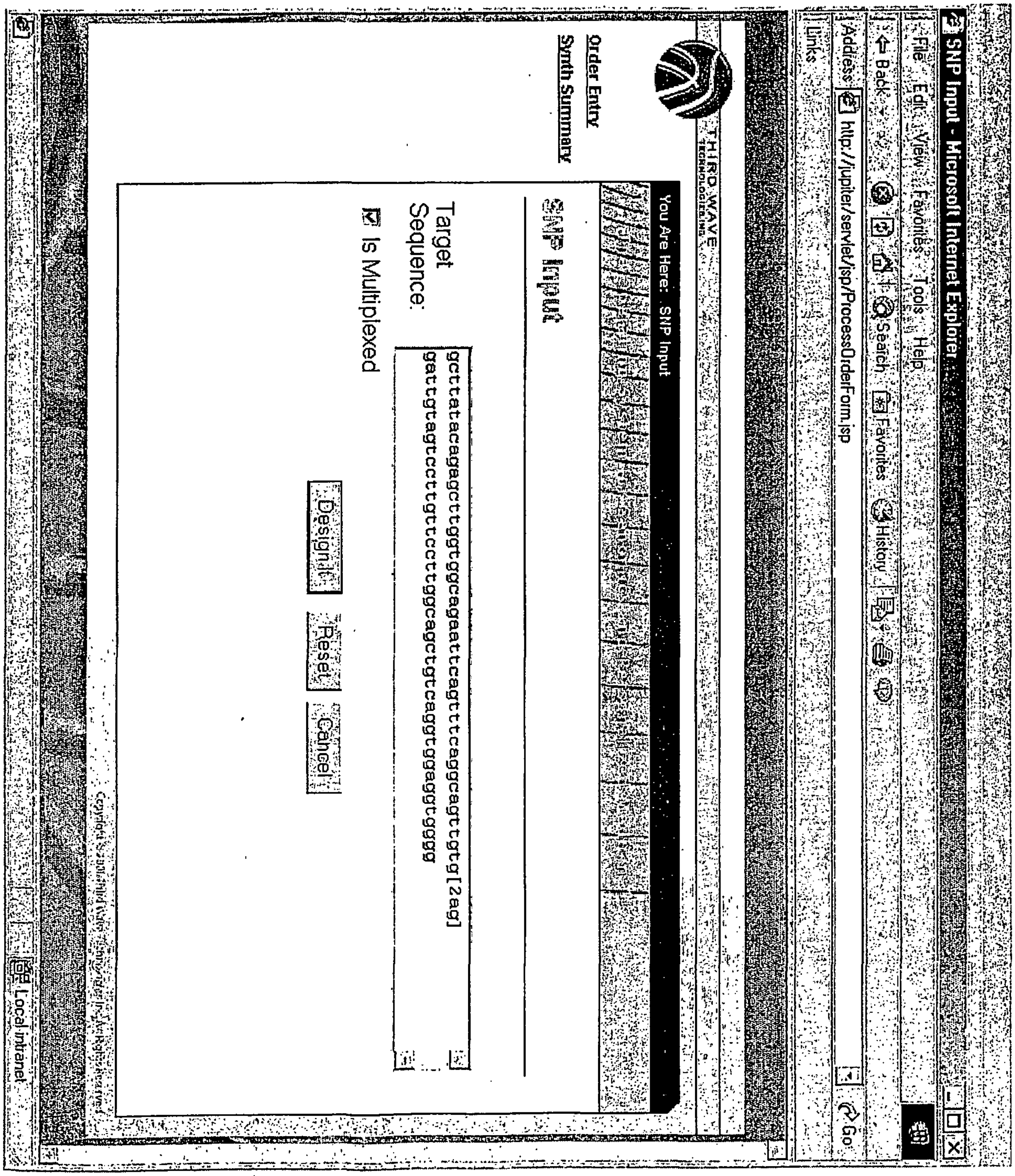

SNP Input - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessOrderForm.jsp
Go
Links
THIRD WAVE
Order Entry
Synth Summary
You Are Here: SNP Input
SNP Input
Target Sequence:
gcttatacaggcttggtggcagaattcagtttcaggcagttgtg[2ag]
gattgtagtccttgttccttggcagctgtcaggtggaggtggg
Is Multiplexed
Design It
Reset
Cancel
Local intranet

FIGURE 12R

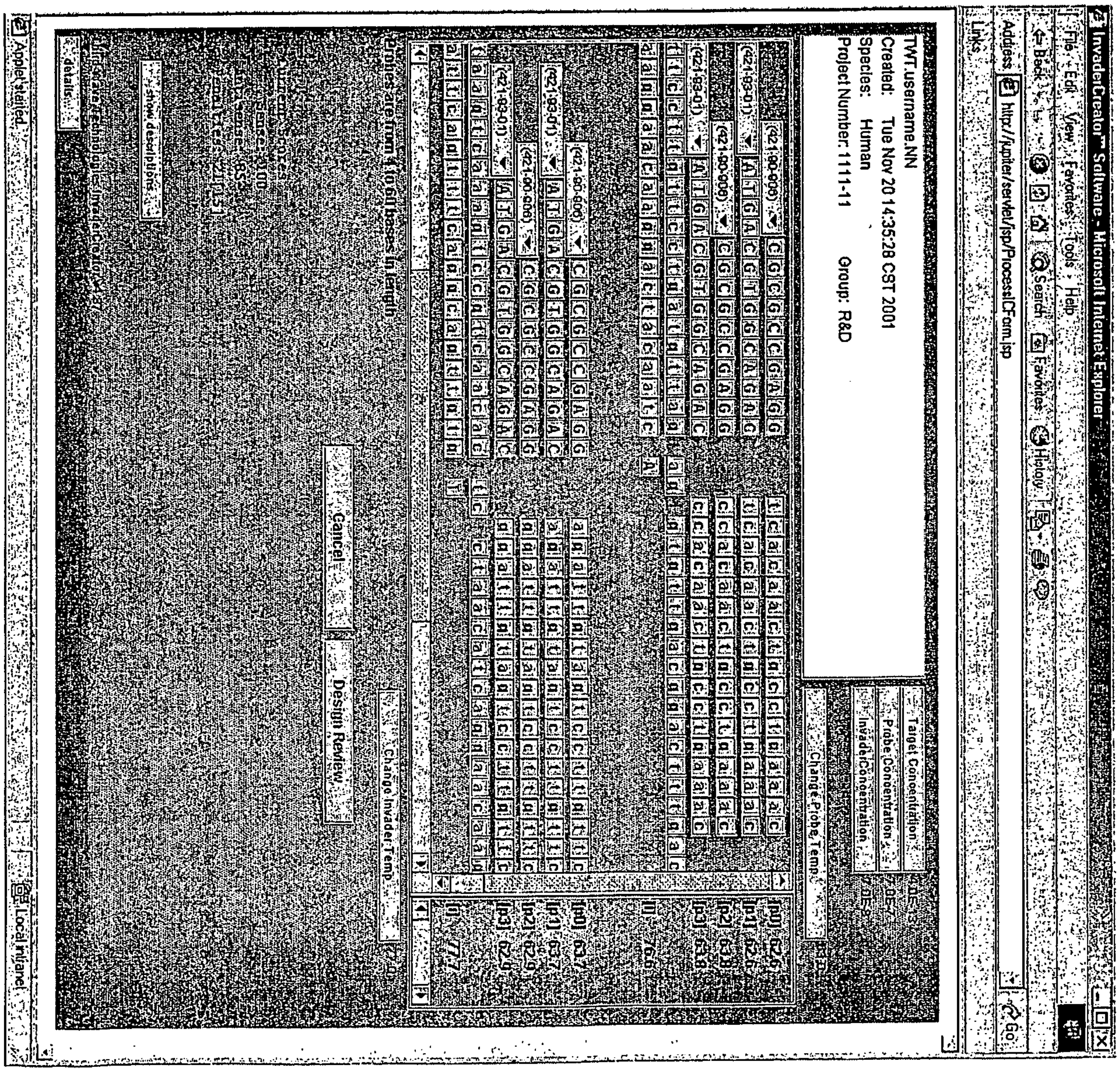
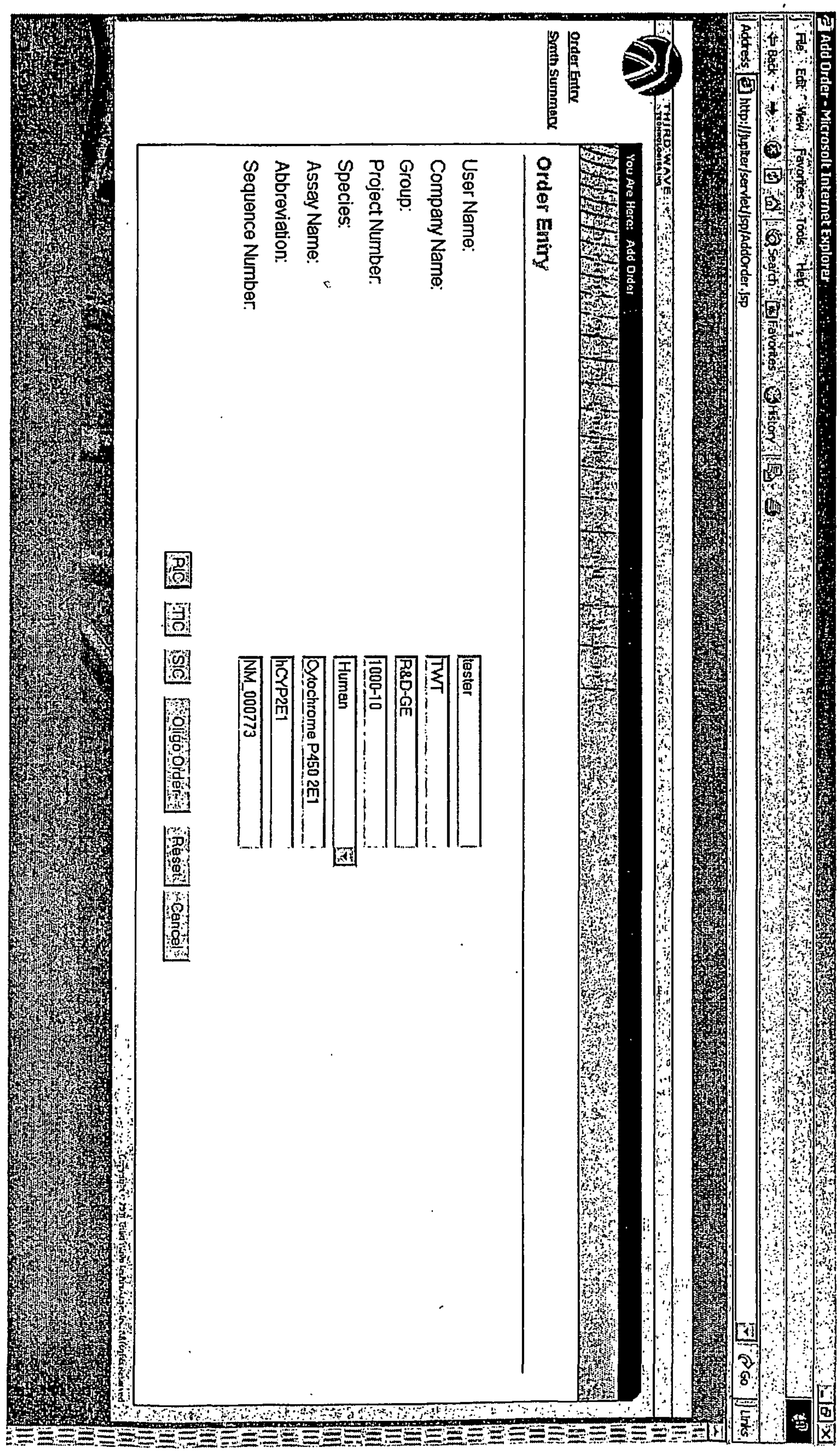

InvaderCreator™ Software - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/ProcessICForm.jsp
Links
Go
TWT.username.NN
Created: Tue Nov 20 14:35:28 CST 2001
Species: Human
Project Number: 1111-11 Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Change Invader Temp
Cancel
Design Review
show descriptions
details...
Applet started
Local intranet

FIGURE 13A


Add Order - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://jupiter/servlet/jsp/AddOrder.jsp
Go Links
THIRD WAVE
Order Entry
Synth Summary
You Are Here: Add Order
Order Entry
User Name: tester
Company Name: TWT
Group: R&D-GE
Project Number: 1000-10
Species: Human
Assay Name: Cytochrome P450 2E1
Abbreviation: hCYP2E1
Sequence Number: NM_000773
PIC
TIC
SIC
Oligo Order
Reset
Cancel

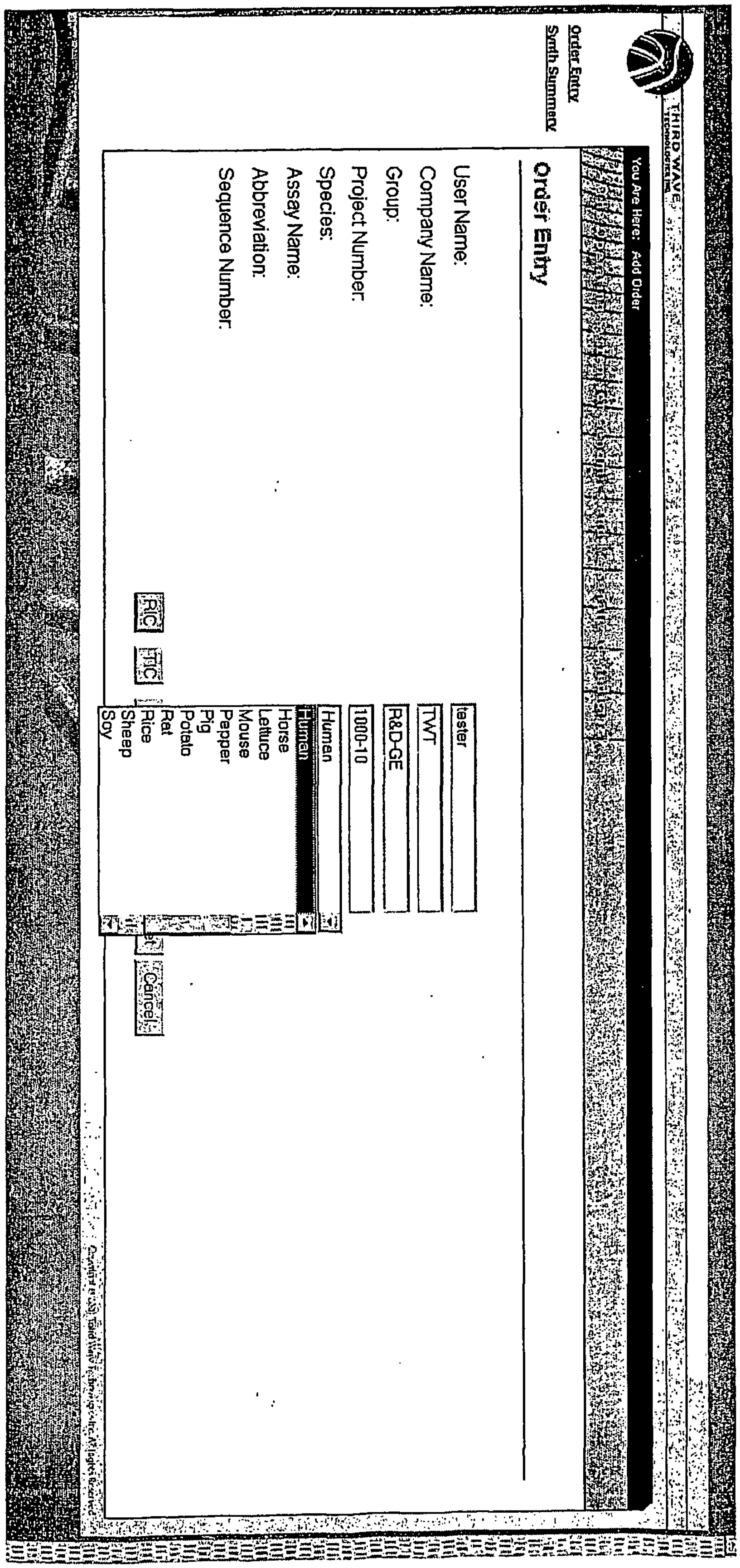
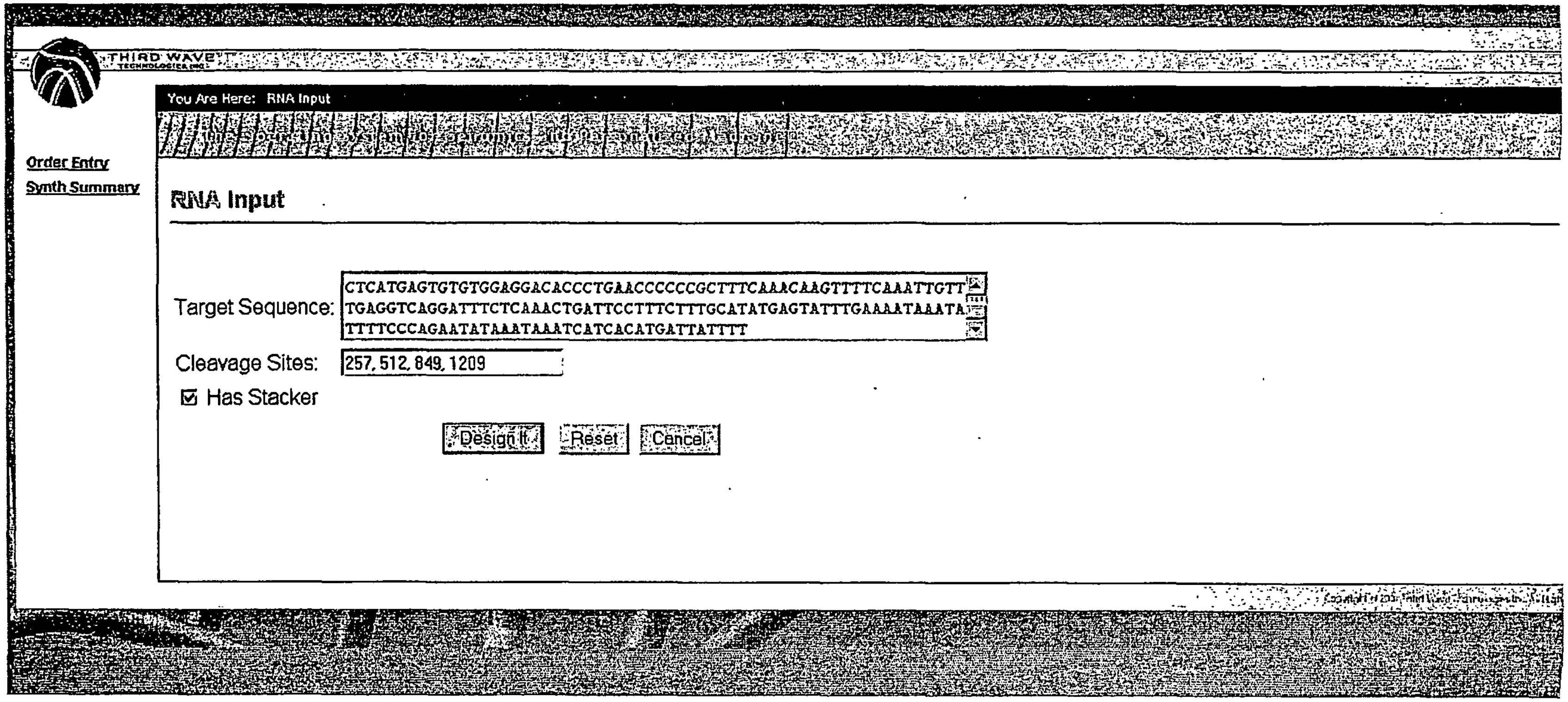

THIRD WAVE
Order Entry
Synth Summary
You Are Here: Add Order
Order Entry
User Name:
tester
Company Name:
TWT
Group:
R&D-GE
Project Number:
1000-10
Species:
Human
Human
Horse
Lettuce
Mouse
Pepper
Pig
Potato
Rat
Rice
Sheep
Soy
Assay Name:
Abbreviation:
Sequence Number:
RIC
FTC
Cancel

FIGURE 13B

FIGURE 13C


THIRD WAVE
You Are Here: RNA Input
Order Entry
Synth Summary
RNA Input
Target Sequence:
CTCATGAGTGTGTGGAGGACACCCTGAACCCCCCGCTTTCAAACAAGTTTTCAAATTGTT
TGAGGTCAGGATTTCTCAAACTGATTCCTTTCTTTGCATATGAGTATTTGAAAATAAATA
TTTTCCCAGAATATAAATAAATCATCACATGATTATTTT
Cleavage Sites:
257, 512, 849, 1209
Has Stacker
Design It
Reset
Cancel

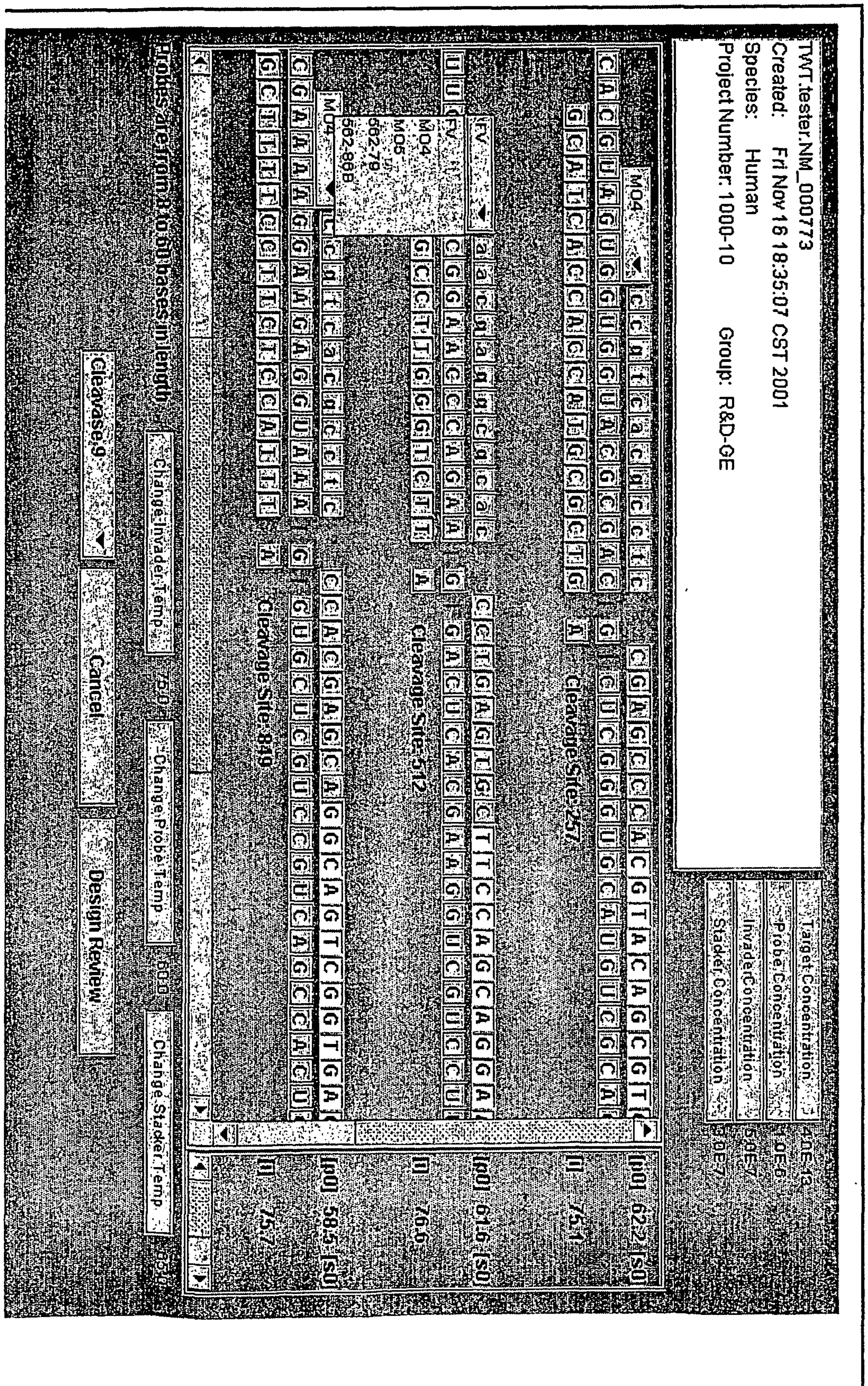

TWT.tester.NM_000773
Created: Fri Nov 16 18:35:07 CST 2001
Species: Human
Project Number: 1000-10 Group: R&D-GE
Cleavage Site 257
Cleavage Site 512
Cleavage Site 849
Target Concentration
Probe Concentration
Invader Concentration
Stacker Concentration
Probes are from 8 to 60 bases in length
Change Invader Temp
Change Probe Temp
Change Stacker Temp
Cleavase 9
Cancel
Design Review

FIGURE 13D

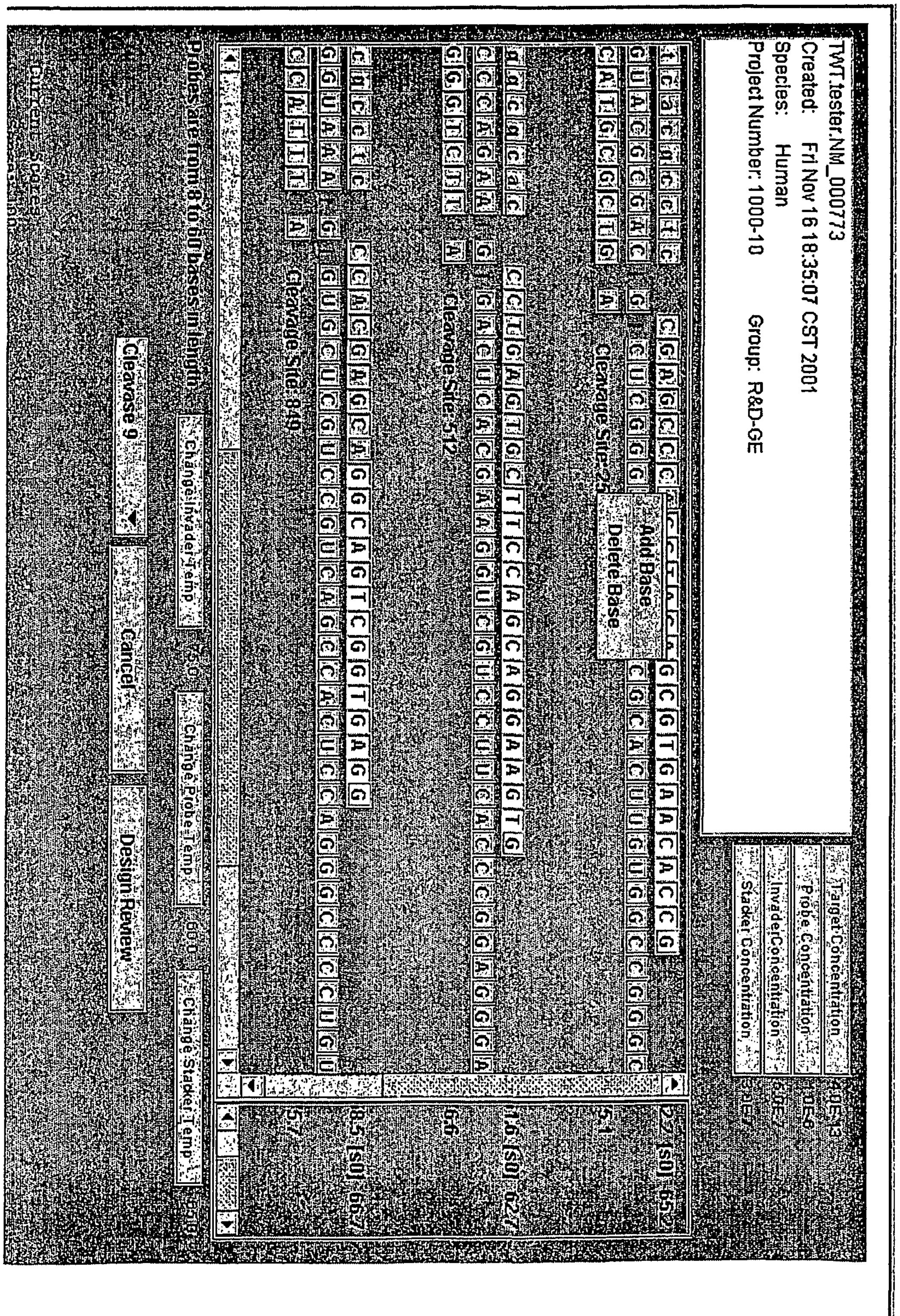
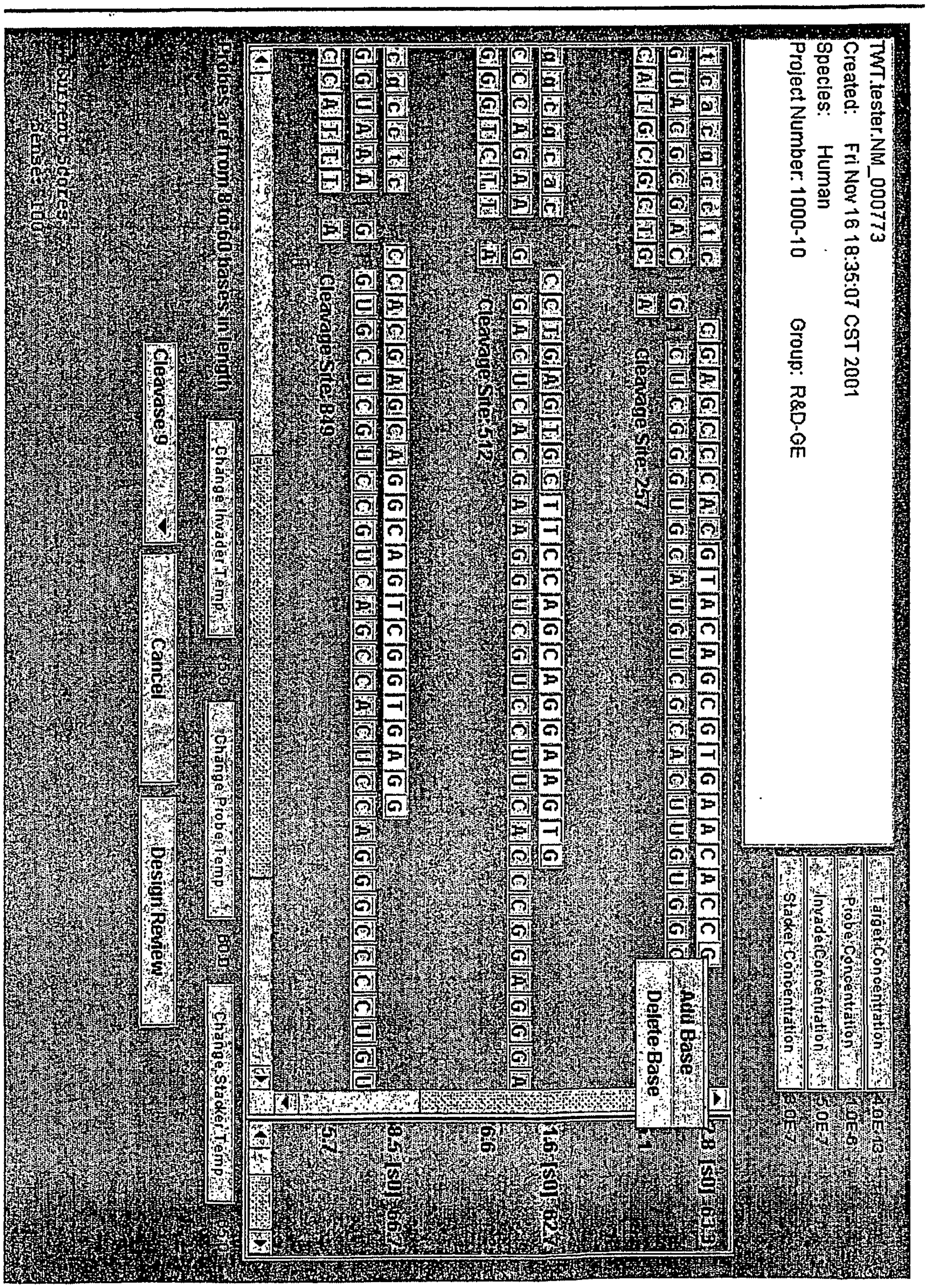

TWT.tester.NM_000773
Created: Fri Nov 16 18:35:07 CST 2001
Species: Human
Project Number: 1000-10 Group: R&D-GE
Target Concentration
Probe Concentration
Invader Concentration
Stacker Concentration
Add Base
Delete Base
Cleavage Site 512
Cleavage Site 849
Probes are from 8 to 60 bases in length
Change Invader Temp
Change Probe Temp
Change Stacker Temp
Cleavase 9
Cancel
Design Review

FIGURE 13E

FIGURE 13F


TWT.tester.NM_000773
Created: Fri Nov 16 18:35:07 CST 2001
Species: Human
Project Number: 1000-10 Group: R&D-GE
Target Concentration
Probe Concentration
Invader Concentration
Stacker Concentration
Add Base
Delete Base
Cleavage Site: 257
Cleavage Site: 512
Cleavage Site: 849
Probes are from 8 to 60 bases in length
Change Invader Temp
Change Probe Temp
Change Stacker Temp
Cleavase 9
Cancel
Design Review
Current Scores

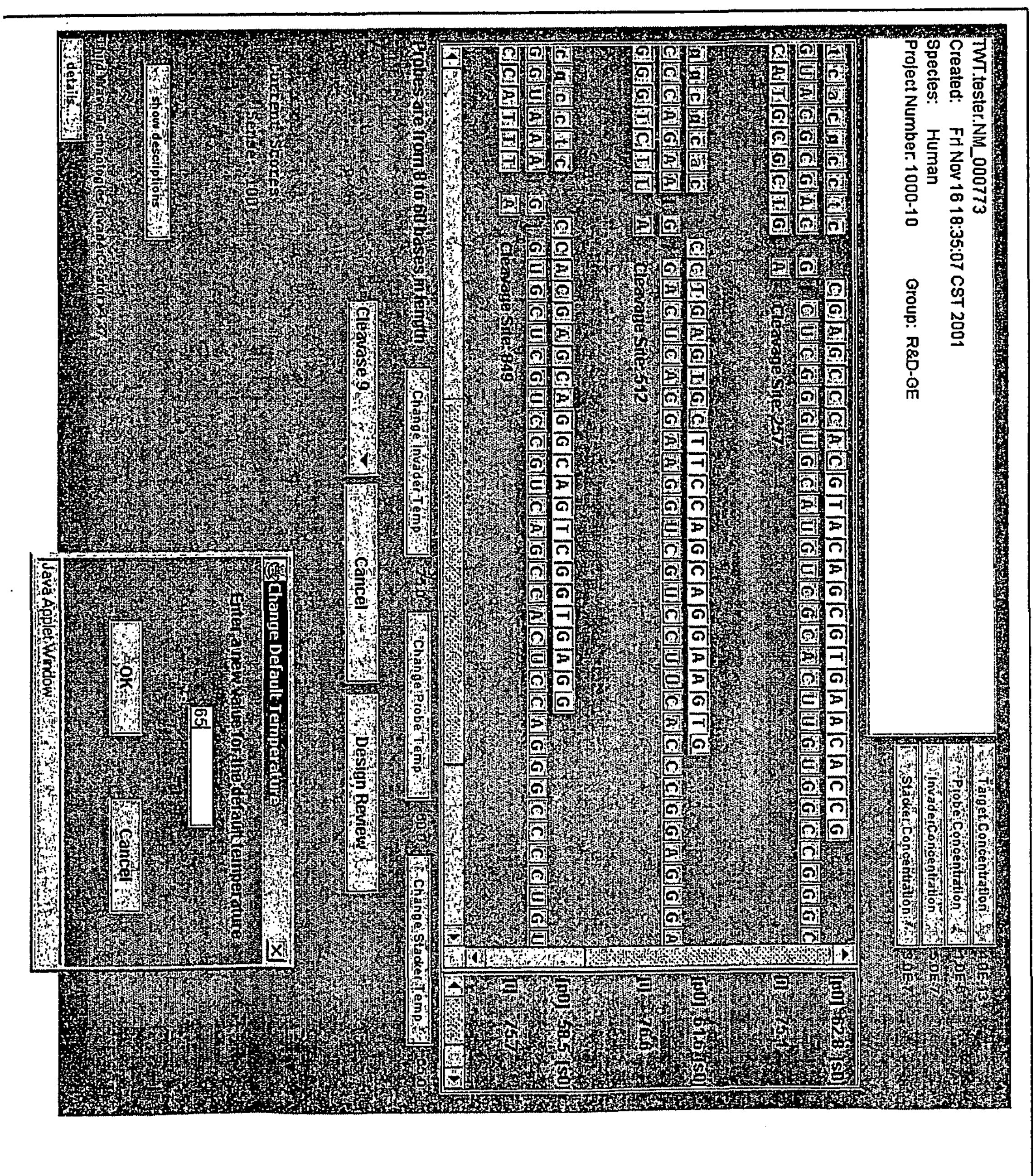

TWT.tester.NM_000773
Created: Fri Nov 16 18:35:07 CST 2001
Species: Human
Project Number: 1000-10
Group: R&D-GE
Target Concentration
Probe Concentration
Invader Concentration
Stacker Concentration
Cleavage Site: 257
Cleavage Site: 512
Cleavage Site: 849
Probes are from 8 to 60 bases in length
Change Invader Temp
Change Probe Temp
Change Stacker Temp
Cleavase 9
Cancel
Design Review
show descriptions
details
Change Default Temperature
Enter a new value for the default temperature
65
OK
Cancel
Java Applet Window

FIGURE 13G

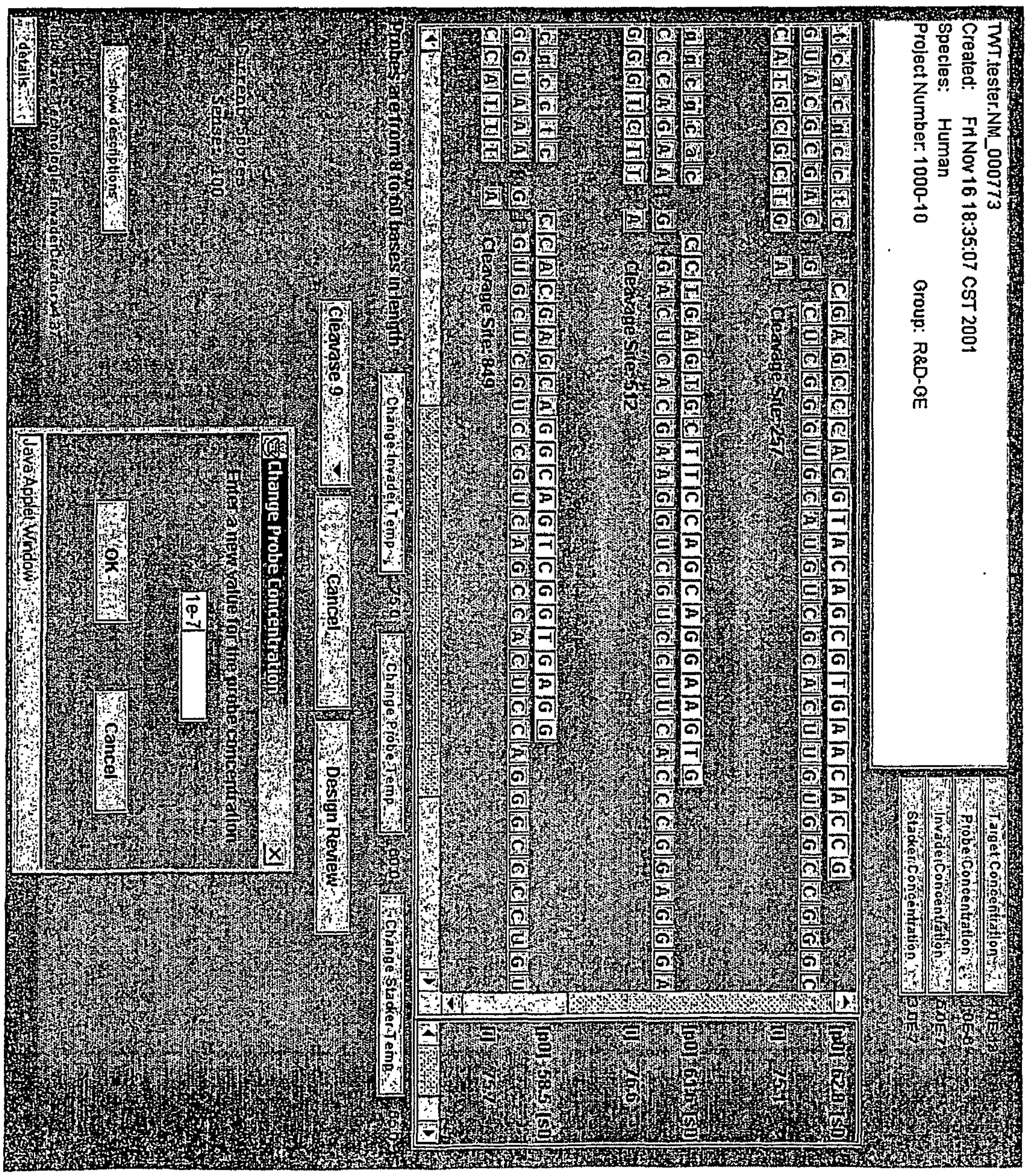

TWT.tester.NM_000773
Created: Fri Nov 16 18:35:07 CST 2001
Species: Human
Project Number: 1000-10 Group: R&D-GE
Cleavage Site: 247
Cleavage Site: 512
Cleavage Site: 849
Target Concentration
Probe Concentration
Invader Concentration
Stacker Concentration
Change Invader Temp.
Change Probe Temp.
Change Stacker Temp.
Cleavase 9
Cancel
Design Review
Change Probe Concentration
Enter a new value for the probe concentration:
1e-7
OK
Cancel
Java Applet Window
show descriptions
details

FIGURE 13H

FIGURE 13I

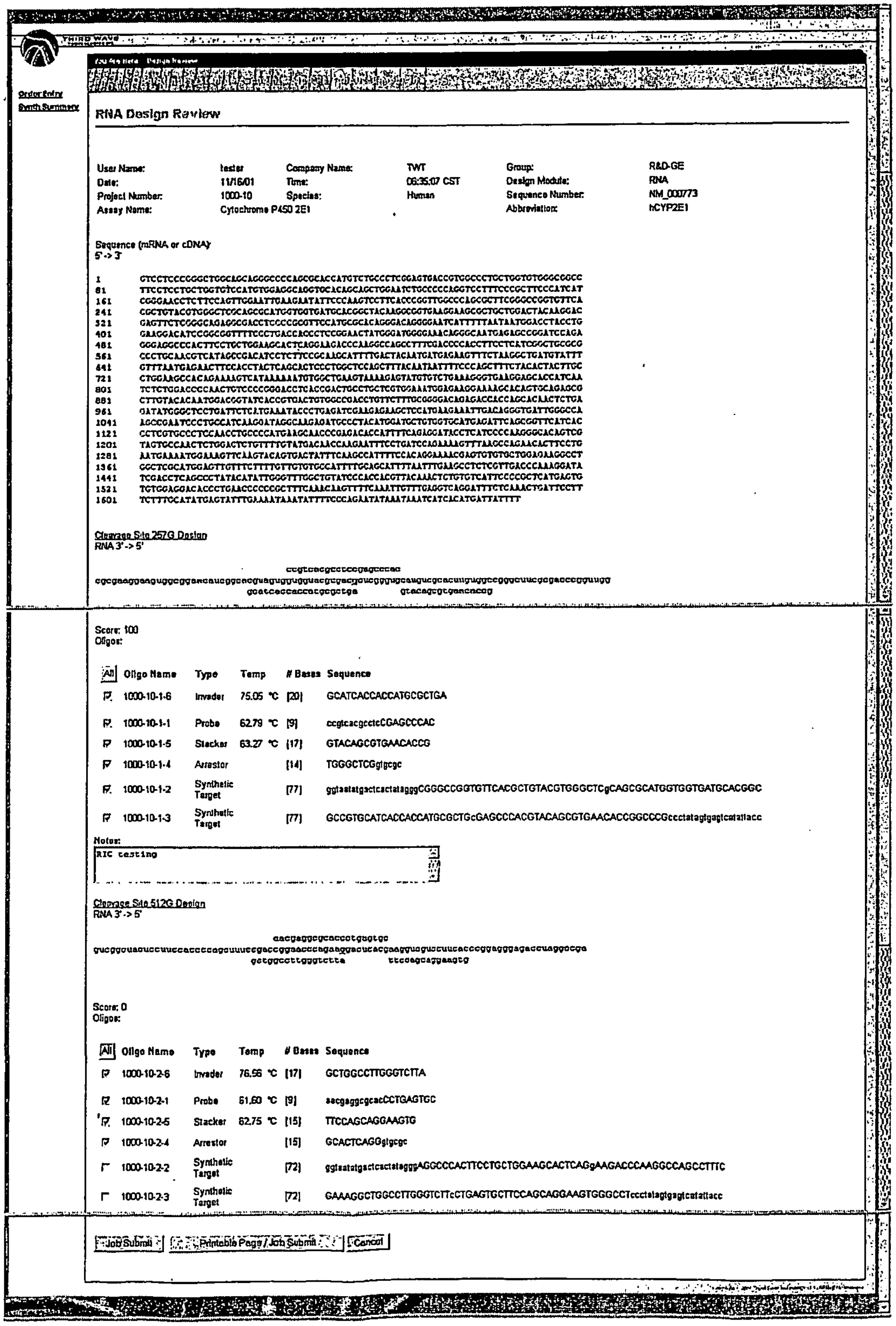

Order Entry
Synth Summary
RNA Design Review
User Name: tester
Company Name: TWT
Group: R&D-GE
Date: 11/16/01
Time: 06:35:07 CST
Design Module: RNA
Project Number: 1000-10
Species: Human
Sequence Number: NM_000773
Assay Name: Cytochrome P450 2E1
Abbreviation: hCYP2E1
Sequence (mRNA or cDNA)
5'-> 3'
Cleavage Site 257G Design
RNA 3'-> 5'
Score: 100
Oligos:
All Oligo Name Type Temp # Bases Sequence
1000-10-1-6 Invader 75.05 °C [20] GCATCACCACCATGCGCTGA
1000-10-1-1 Probe 62.79 °C [9] ccgtcacgcctcCGAGCCCAC
1000-10-1-5 Stacker 63.27 °C [17] GTACAGCGTGAACACCG
1000-10-1-4 Arrestor [14] TGGGCTCGgtgcgc
1000-10-1-2 Synthetic Target [77] ggtaatatgactcactatagggCGGGCCGGTGTTCACGCTGTACGTGGGCTCgCAGCGCATGGTGGTGATGCACGGC
1000-10-1-3 Synthetic Target [77] GCCGTGCATCACCACCATGCGCTGcGAGCCCACGTACAGCGTGAACACCGGCCCGccctatagtgagtcatattacc
Notes:
RIC testing
Cleavage Site 512G Design
RNA 3'-> 5'
Score: 0
Oligos:
All Oligo Name Type Temp # Bases Sequence
1000-10-2-6 Invader 76.56 °C [17] GCTGGCCTTGGGTCTTA
1000-10-2-1 Probe 61.60 °C [9] aacgaggcgcacCCTGAGTGC
1000-10-2-5 Stacker 62.75 °C [15] TTCCAGCAGGAAGTG
1000-10-2-4 Arrestor [15] GCACTCAGGgtgcgc
1000-10-2-2 Synthetic Target [72] ggtaatatgactcactatagggAGGCCCACTTCCTGCTGGAAGCACTCAGgAAGACCCAAGGCCAGCCTTTC
1000-10-2-3 Synthetic Target [72] GAAAGGCTGGCCTTGGGTCTTcCTGAGTGCTTCCAGCAGGAAGTGGGCCTccctatagtgagtcatattacc
Job Submit
Printable Page / Job Submit
Cancel

FIGURE 13J

User Name:	tester	Company Name:	TWT	Group:	R&D-GE
Date:	11/16/01	Time:	06:35:07 CST	Design Module:	RNA
Project Number:	1000-10	Species:	Human	Sequence Number:	NM_000773
Assay Name:	Cytochrome P450 2E1			Abbreviation:	hCYP2E1

Sequence (mRNA or cDNA):
5' -> 3'

```
1     GTCCTCCCGGGCTGGCAGCAGGGCCCCAGCGCACCATGTCTGCCCTCGGAGTGACCGTGGCCCTGCTGGTGTGGGCGGCC
81    TTCCTCCTGCTGGTGTCCATGTGGAGGCAGGTGCACAGCAGCTGGAATCTGCCCCCAGGTCCTTTCCCGCTTCCCATCAT
161   CGGGAACCTCTTCCAGTTGGAATTGAAGAATATTCCCAAGTCCTTCACCCGGTTGGCCCAGCGCTTCGGGCCGGTGTTCA
241   CGCTGTACGTGGGCTCGCAGCGCATGGTGGTGATGCACGGCTACAAGGCGGTGAAGGAAGCGCTGCTGGACTACAAGGAC
321   GAGTTCTCGGGCAGAGGCGACCTCCCCGCGTTCCATGCGCACAGGGACAGGGGAATCATTTTTAATAATGGACCTACCTG
401   GAAGGACATCCGGCGGTTTTCCCTGACCACCCTCCGGAACTATGGGATGGGGAAACAGGGCAATGAGAGCCGGATCCAGA
481   GGGAGGCCCACTTCCTGCTGGAAGCACTCAGGAAGACCCAAGGCCAGCCTTTCGACCCCACCTTCCTCATCGGCTGCGCG
561   CCCTGCAACGTCATAGCCGACATCCTCTTCCGCAAGCATTTTGACTACAATGATGAGAAGTTTCTAAGGCTGATGTATTT
641   GTTTAATGAGAACTTCCACCTACTCAGCACTCCCTGGCTCCAGCTTTACAATAATTTTCCCAGCTTTCTACACTACTTGC
721   CTGGAAGCCACAGAAAAGTCATAAAAAATGTGGCTGAAGTAAAAGAGTATGTGTCTGAAAGGGTGAAGGAGCACCATCAA
801   TCTCTGGACCCCAACTGTCCCCGGGACCTCACCGACTGCCTGCTCGTGGAAATGGAGAAGGAAAAGCACAGTGCAGAGCG
881   CTTGTACACAATGGACGGTATCACCGTGACTGTGGCCGACCTGTTCTTTGCGGGGACAGAGACCACCAGCACAACTCTGA
961   GATATGGGCTCCTGATTCTCATGAAATACCCTGAGATCGAAGAGAAGCTCCATGAAGAAATTGACAGGGTGATTGGGCCA
1041  AGCCGAATCCCTGCCATCAAGGATAGGCAAGAGATGCCCTACATGGATGCTGTGGTGCATGAGATTCAGCGGTTCATCAC
1121  CCTCGTGCCCTCCAACCTGCCCCATGAAGCAACCCGAGACACCATTTTCAGAGGATACCTCATCCCCAAGGGCACAGTCG
1201  TAGTGCCAACTCTGGACTCTGTTTTGTATGACAACCAAGAATTTCCTGATCCAGAAAAGTTTAAGCCAGAACACTTCCTG
1281  AATGAAAATGGAAAGTTCAAGTACAGTGACTATTTCAAGCCATTTTCCACAGGAAAACGAGTGTGTGCTGGAGAAGGCCT
1361  GGCTCGCATGGAGTTGTTTCTTTTGTTGTGTGCCATTTTGCAGCATTTTAATTTGAAGCCTCTCGTTGACCCAAAGGATA
1441  TCGACCTCAGCCCTATACATATTGGGTTTGGCTGTATCCCACCACGTTACAAACTCTGTGTCATTCCCCGCTCATGAGTG
1521  TGTGGAGGACACCCTGAACCCCCCGCTTTCAAACAAGTTTTCAAATTGTTTGAGGTCAGGATTTCTCAAACTGATTCCTT
1601  TCTTTGCATATGAGTATTTGAAAATAAATATTTTCCCAGAATATAAATAAATCATCACATGATTATTTT
```

Cleavage Site 257G Design
RNA 3' -> 5'

```
                                                     ccgtcacgcctccgagcccac
cgcgaaggaaguggcggaacaucggcacguaguggugguacgcgacgcucgggugcaugucgcacuuguggccgggcuucgcgacccgguugg
                                               gcatcaccaccatgcgctga            gtacagcgtgaacaccg
```

Score: 100
Oligos:

	Oligo Name	Type	Temp	# Bases	Sequence
•	1000-10-1-6	Invader	75.05 °C	[20]	GCATCACCACCATGCGCTGA
•	1000-10-1-1	Probe	62.79 °C	[9]	ccgtcacgcctcCGAGCCCAC
•	1000-10-1-5	Stacker	63.27 °C	[17]	GTACAGCGTGAACACCG
•	1000-10-1-4	Arrestor		[14]	TGGGCTCGgtgcgc
•	1000-10-1-2	Synthetic Target		[77]	ggtaatatgactcactatagggCGGGCCGGTGTTCACGCTGTACGTGGGCTCgCAGCGCATGGTGGTGATGCACGGC
•	1000-10-1-3	Synthetic Target		[77]	GCCGTGCATCACCACCATGCGCTGcGAGCCCACGTACAGCGTGAACACCGGCCCGccctatagtgagtcatattacc

Notes:
RIC testing

Cleavage Site 512G Design
RNA 3' -> 5'

```
                                                   aacgaggcgcaccctgagtgc
gucggcuacuccuuccaccccagcuuuccgaccggaacccagaaggacucacgaaggucguccuucacccggagggagaccuaggccga
                                             gctggccttgggtctta            ttccagcaggaagtg
```

Score: 0
Oligos:

	Oligo Name	Type	Temp	# Bases	Sequence
•	1000-10-2-6	Invader	76.56 °C	[17]	GCTGGCCTTGGGTCTTA
•	1000-10-2-1	Probe	61.60 °C	[9]	aacgaggcgcacCCTGAGTGC
•	1000-10-2-5	Stacker	62.75 °C	[15]	TTCCAGCAGGAAGTG
•	1000-10-2-4	Arrestor		[15]	GCACTCAGGgtgcgc
	1000-10-2-2	Synthetic Target		[72]	ggtaatatgactcactatagggAGGCCCACTTCCTGCTGGAAGCACTCAGgAAGACCCAAGGCCAGCCTTTC
	1000-10-2-3	Synthetic Target		[72]	GAAAGGCTGGCCTTGGGTCTTcCTGAGTGCTTCCAGCAGGAAGTGGGCCTccctatagtgagtcatattacc

Notes:

FIGURE 13K

THIRD WAVE TECHNOLOGIES INC.

You Are Here: Oligos Ready For Job Submit

Order Entry
Synth Summary

Oligos Ready For Job Submit

Order Number: 24
Order Type: RNA
Date Ordered: 11/16/01
Requester Name: tester

Oligo Name	Oligo Type	Sequence
1000-10-1-1	Probe	ccgtcacgcctcCGAGCCCAC
1000-10-2-1	Probe	aacgaggcgcacCCTGAGTGC
1000-10-1-6	Invader	GCATCACCACCATGCGCTGA
1000-10-2-6	Invader	GCTGGCCTTGGGTCTTA
1000-10-1-4	Arrestor	TGGGCTCGgtgcgc
1000-10-2-4	Arrestor	GCACTCAGGgtgcgc
1000-10-1-5	Stacker	GTACAGCGTGAACACCG
1000-10-2-5	Stacker	TTCCAGCAGGAAGTG
1000-10-1-2	Synthetic Target	ggtaatatgactcactatagggCGGGCCGGTGTTCACGCTGTACGTGGGCTCgCAGCGCATGGTGGTGATGCACGGC
1000-10-1-3	Synthetic Target	GCCGTGCATCACCACCATGCGCTGcGAGCCCACGTACAGCGTGAACACCGGCCCGccctatagtgagtcatattacc

FIGURE 13L

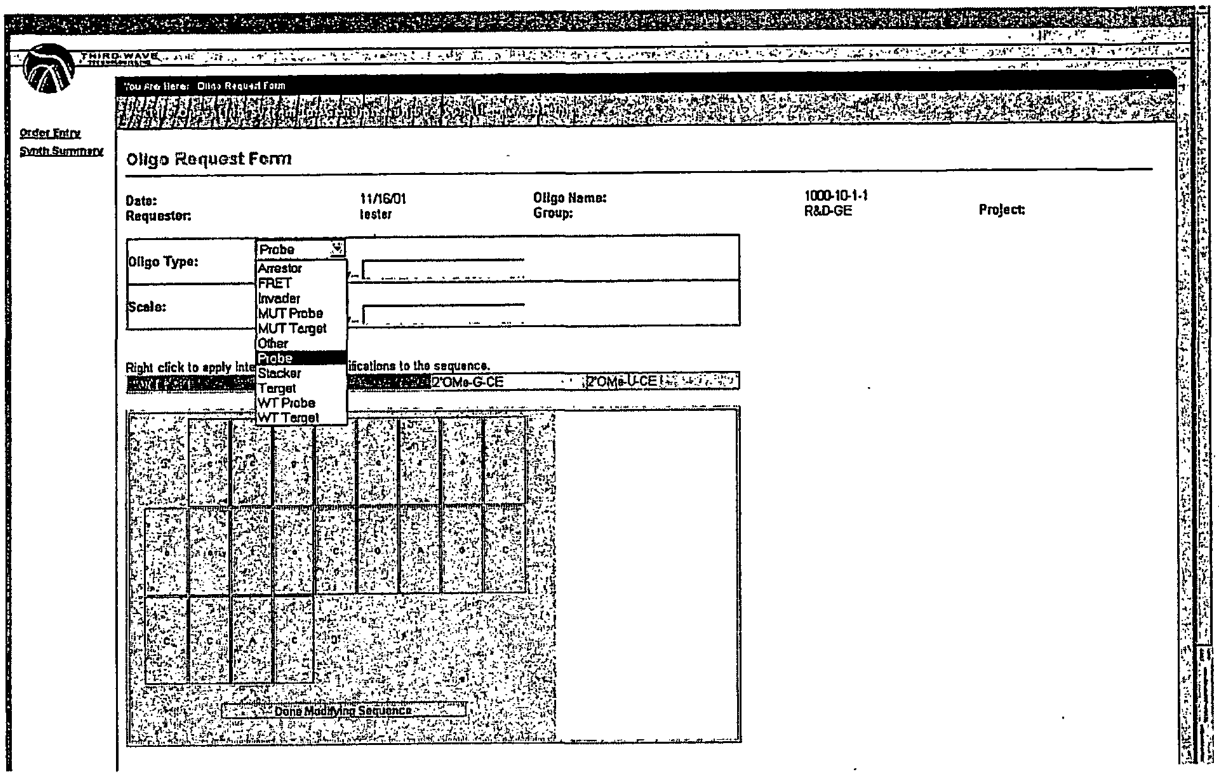


FIGURE 13M

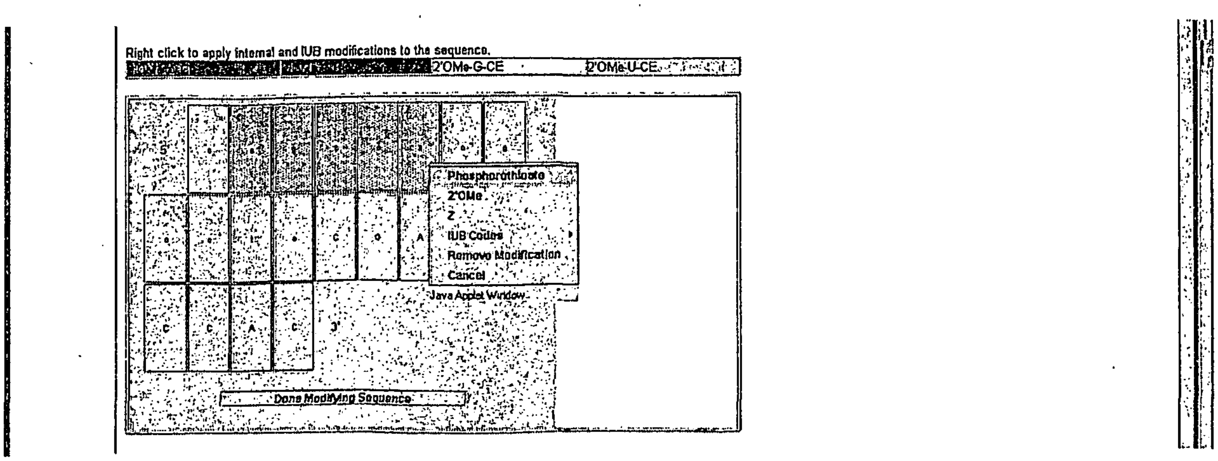


FIGURE 13N

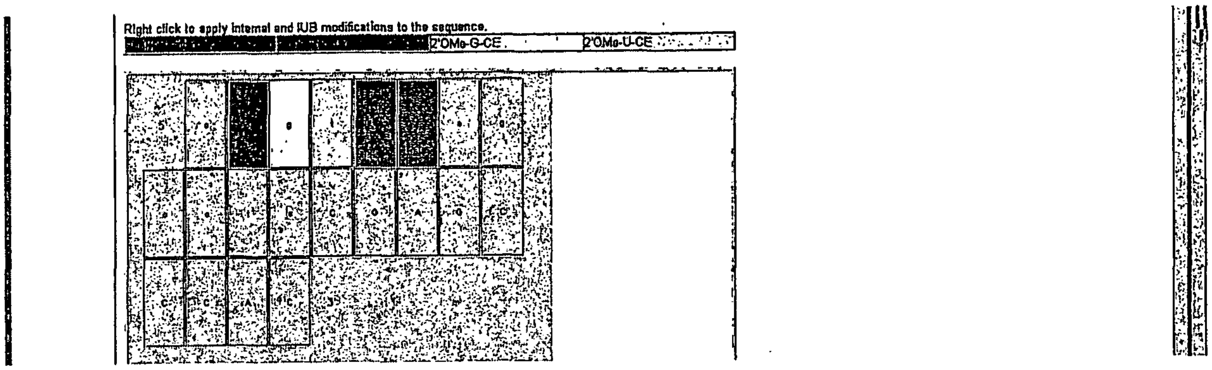
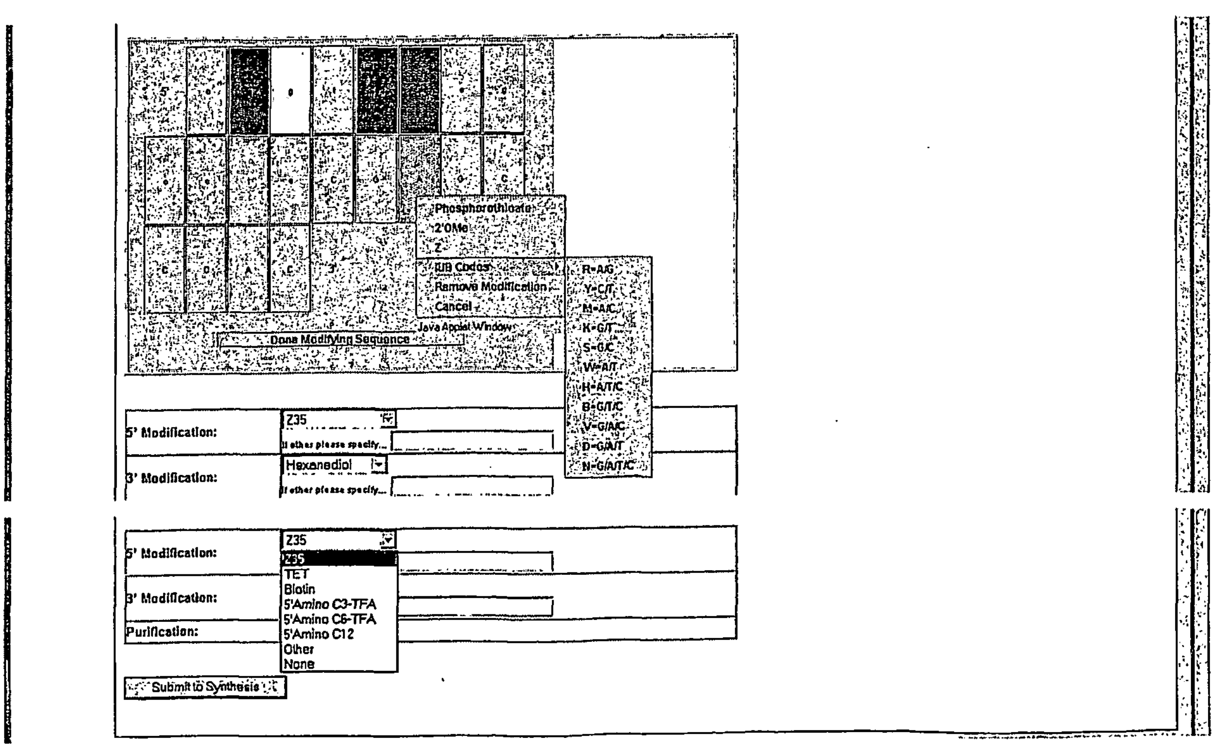


FIGURE 13O



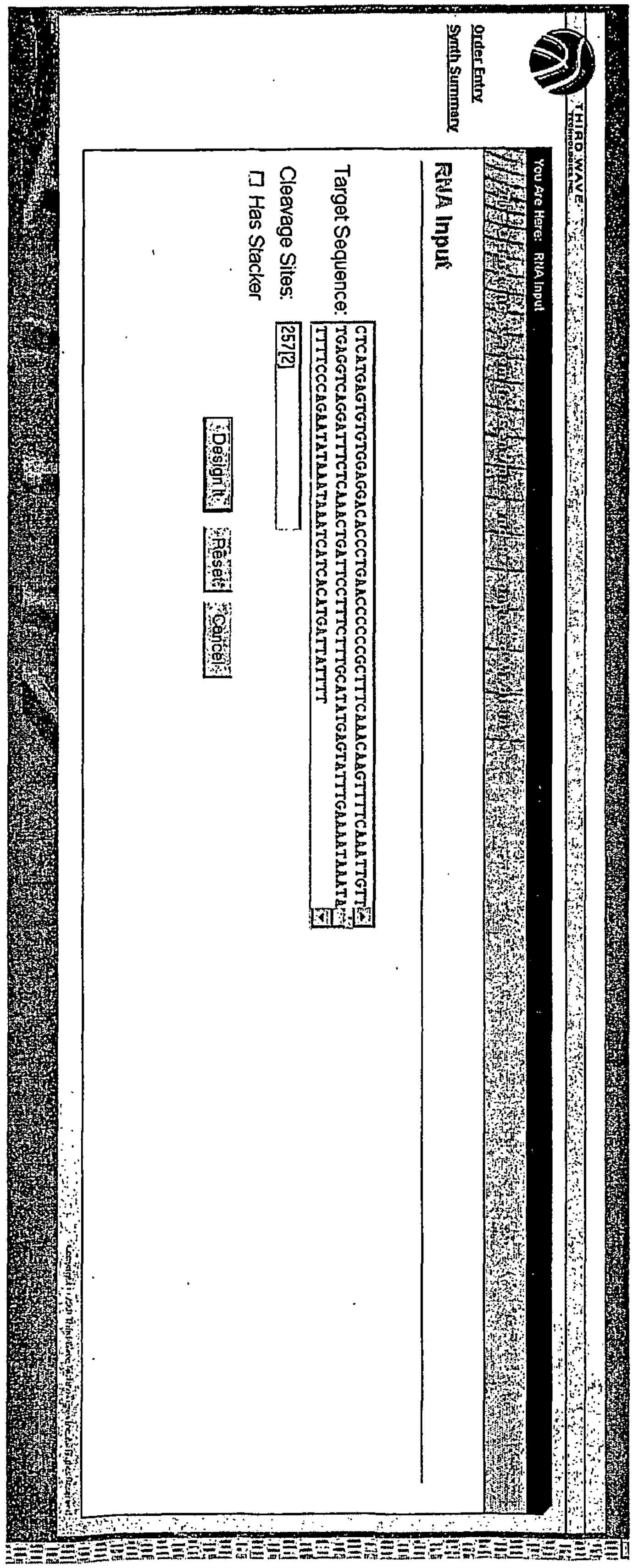

THIRD WAVE
Order Entry
Synth Summary
You Are Here: RNA Input
RNA Input
Target Sequence:
Cleavage Sites:
257(2)
Has Stacker
Design It
Reset
Cancel

FIGURE 13P

FIGURE 13Q

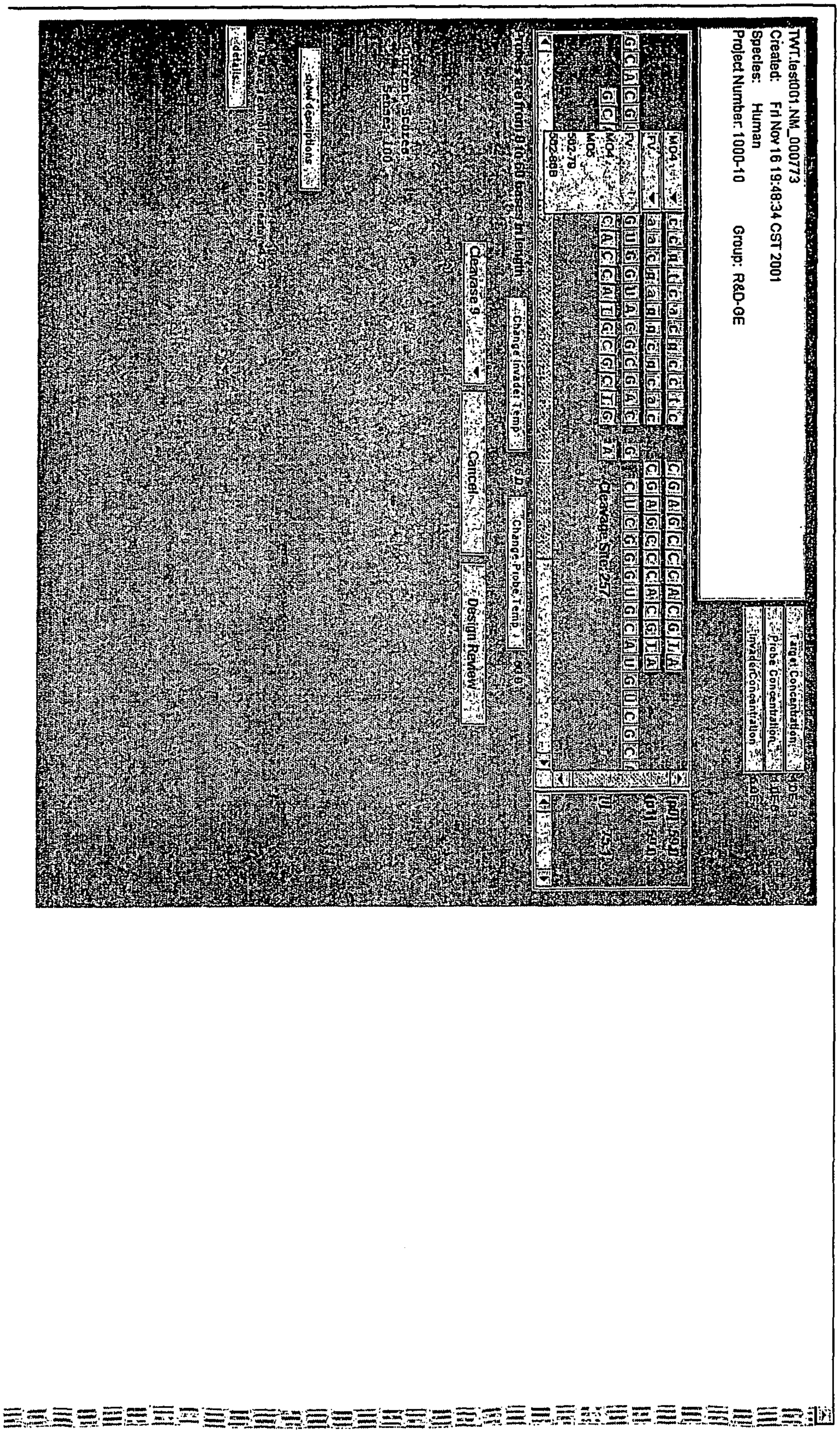

TWT.test001.NM_000773
Created: Fri Nov 16 19:48:34 CST 2001
Species: Human
Project Number: 1000-10 Group: R&D-GE
Target Concentration
Probe Concentration
Invader Concentration
Change Invader Temp
Change Probe Temp
Cleavase 9
Cancel
Design Review
show descriptions
details...

FIGURE 14A

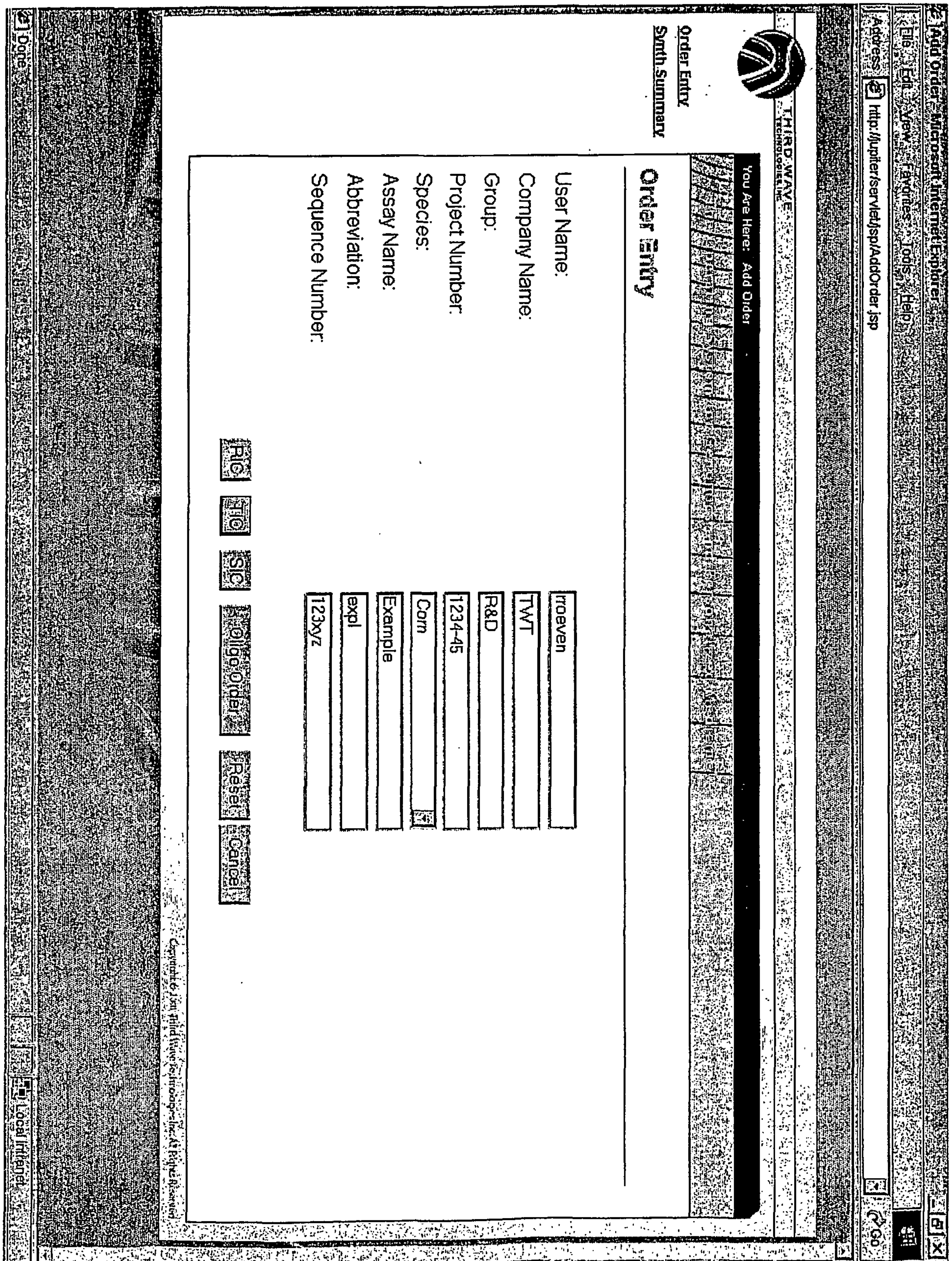

Add Order - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://jupiter/servlet/jsp/AddOrder.jsp
Go
THIRD WAVE TECHNOLOGIES INC
You Are Here: Add Order
Order Entry
Synth Summary
Order Entry
User Name:
Company Name:
Group:
Project Number:
Species:
Assay Name:
Abbreviation:
Sequence Number:
rroeven
TWT
R&D
1234-45
Corn
Example
expl
123xyz
RIC
TIC
SIC
Oligo Order
Reset
Cancel
Done
Local intranet

FIGURE 14B

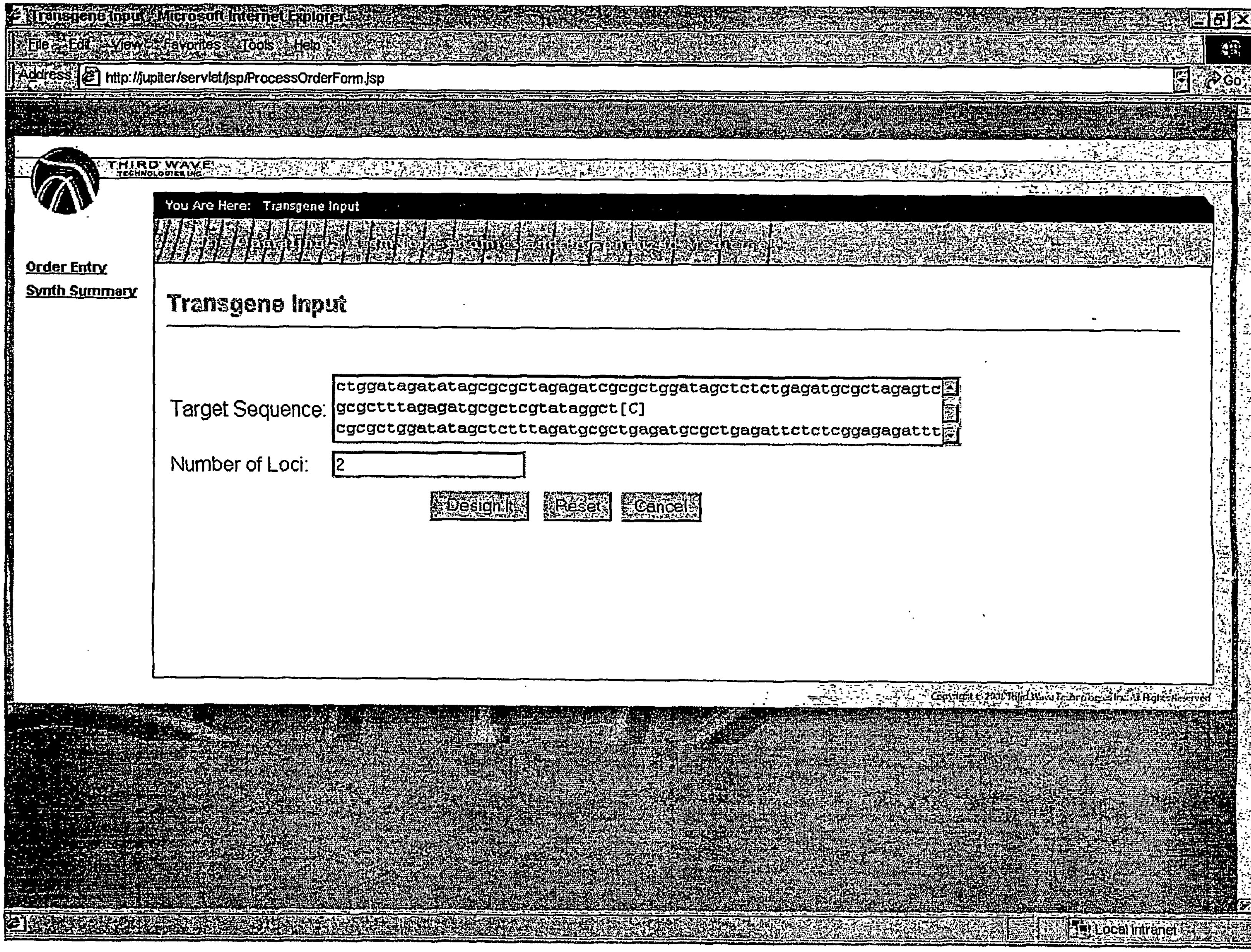

Transgene Input - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://jupiter/servlet/jsp/ProcessOrderForm.jsp
Go
THIRD WAVE
You Are Here: Transgene Input
Order Entry
Synth Summary
Transgene Input
Target Sequence:
ctggatagatatagcgcgctagagatcgcgctggatagctctctgagatgcgctagagtc
gcgctttagagatgcgctcgtataggct[C]
cgcgctggatatagctctttagatgcgctgagatgcgctgagattctctcggagagattt
Number of Loci:
2
Design It
Reset
Cancel
Local intranet

FIGURE 14C

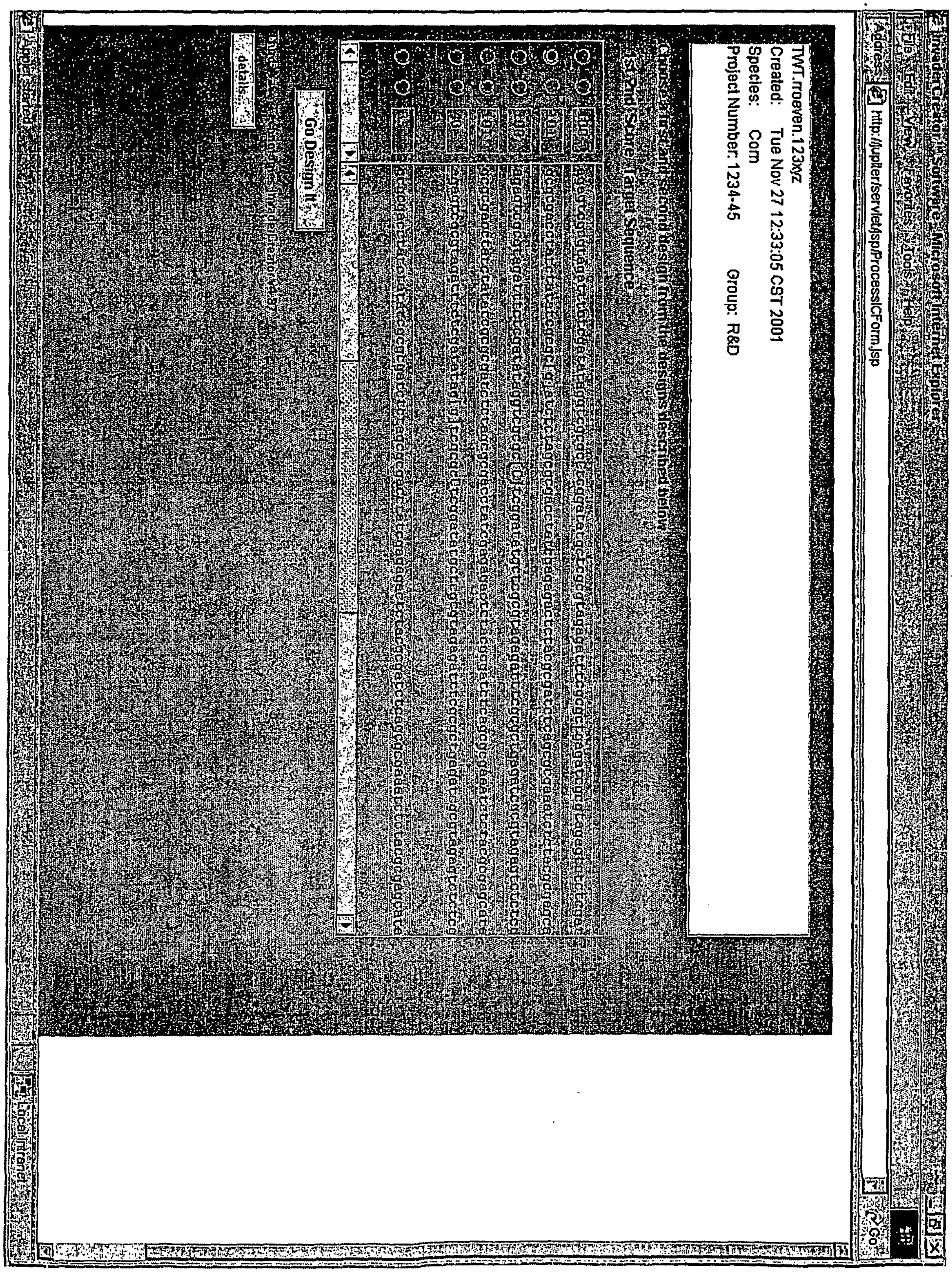

http://jupiter/servlet/jspProcessICForm.jsp
TWT.rroeven.123xyz
Created: Tue Nov 27 12:33:05 CST 2001
Species: Corn
Project Number: 1234-45
Group: R&D
Go Design It
details

FIGURE 14D

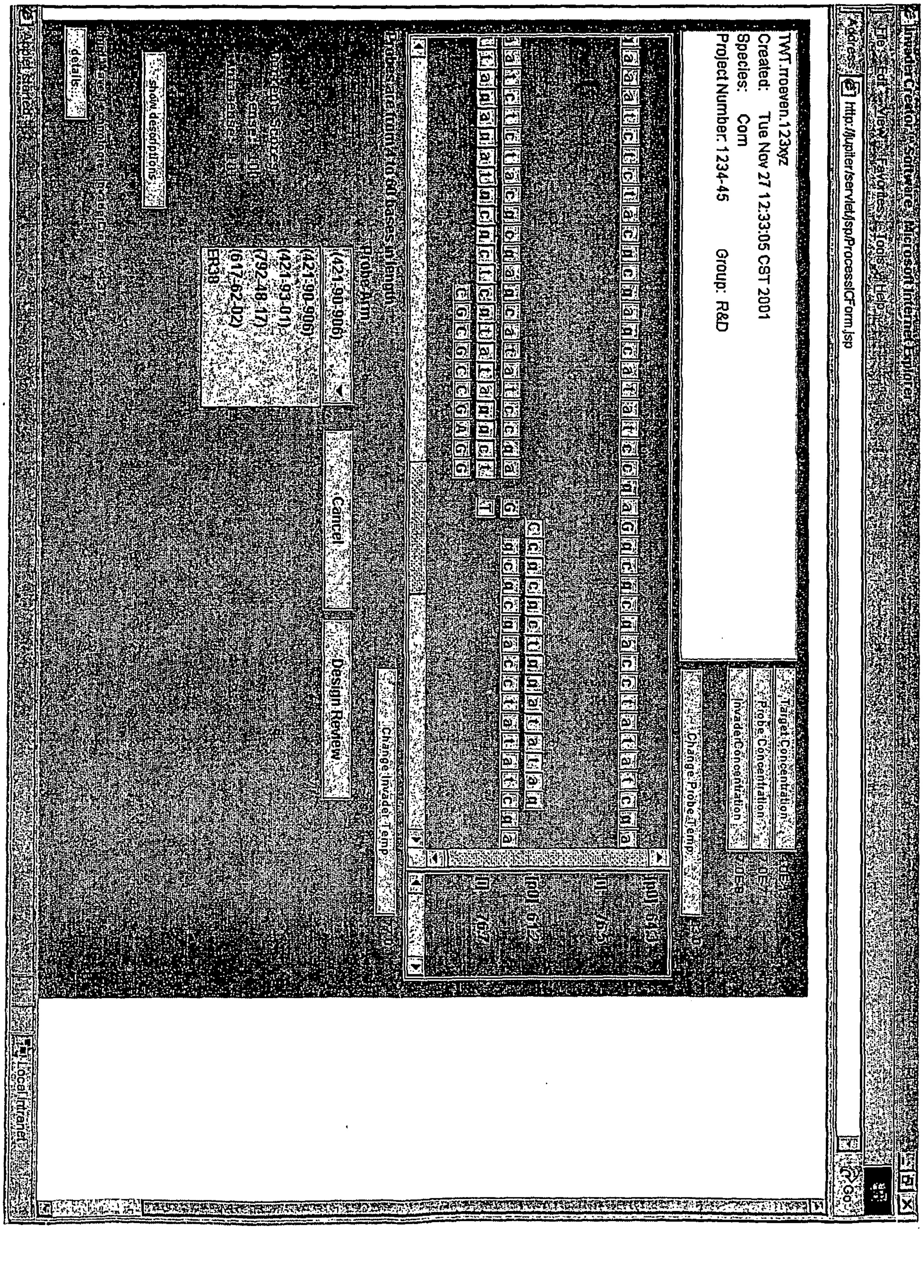

Invader Creator Software - Microsoft Internet Explorer
http://jupiter/servlet/jspProcessICForm.jsp
TWT.rroeven.123xyz
Created: Tue Nov 27 12:33:05 CST 2001
Species: Corn
Project Number: 1234-45 Group: R&D
Target Concentration
Probe Concentration
Invader Concentration
Change Probe Temp
Change Invader Temp
Probe Arm
(421-90-906)
(421-93-01)
(792-48-17)
(617-62-02)
ER38
Cancel
Design Review
show descriptions
details
Applet started
Local intranet

FIGURE 14E

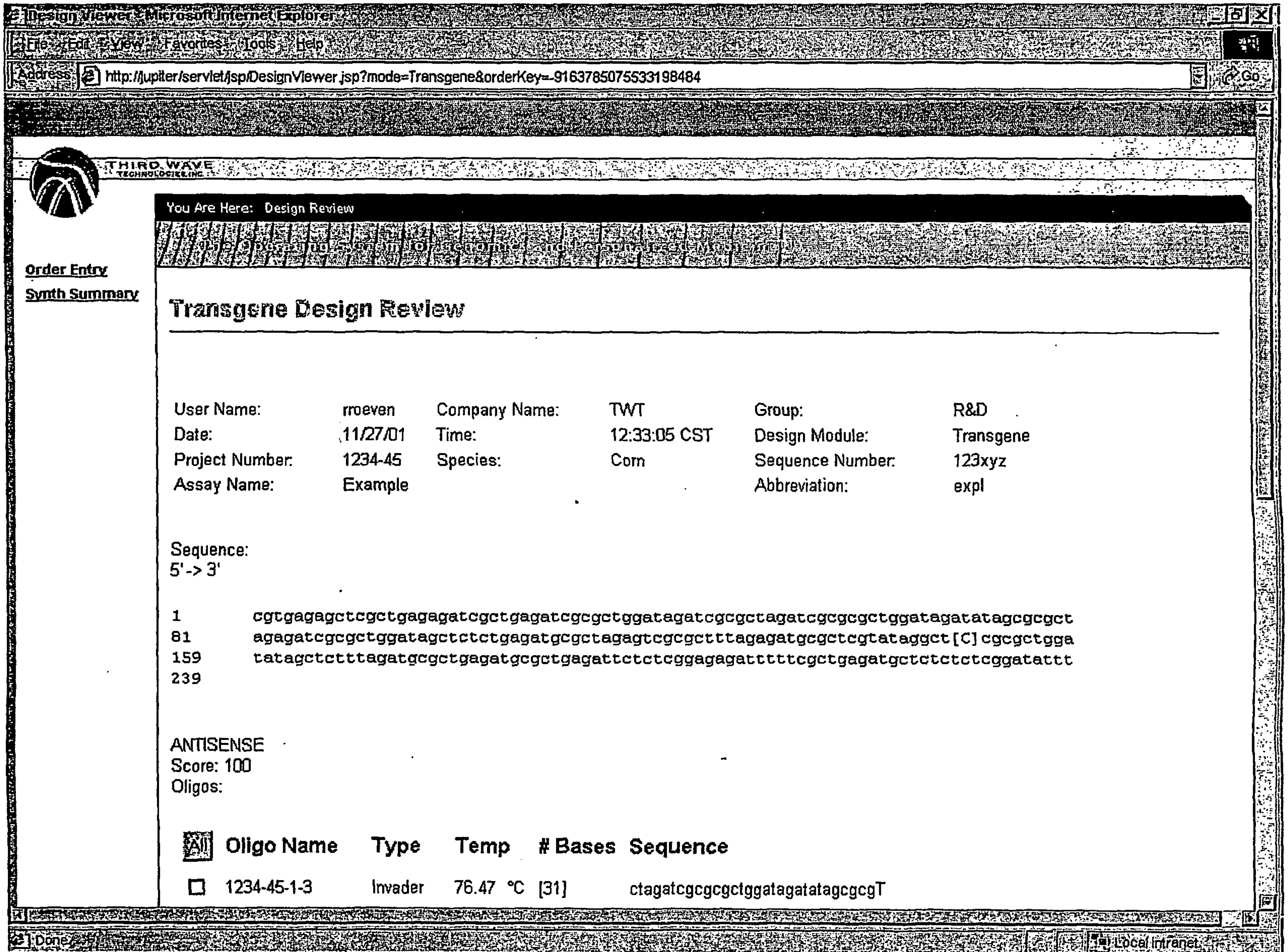

Design Viewer - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://jupiter/servlet/jsp/DesignViewer.jsp?mode=Transgene&orderKey=-9163785075533198484
Go
THIRD WAVE TECHNOLOGIES INC
You Are Here: Design Review
Order Entry
Synth Summary
Transgene Design Review
User Name: rroeven
Company Name: TWT
Group: R&D
Date: 11/27/01
Time: 12:33:05 CST
Design Module: Transgene
Project Number: 1234-45
Species: Corn
Sequence Number: 123xyz
Assay Name: Example
Abbreviation: expl
Sequence:
5' -> 3'
1 cgtgagagctcgctgagagatcgctgagatcgcgctggatagatcgcgctagatcgcgcgctggatagatatagcgcgct
81 agagatcgcgctggatagctctctgagatgcgctagagtcgcgctttagagatgcgctcgtataggct[C]cgcgctgga
159 tatagctctttagatgcgctgagatgcgctgagattctctcggagagattttcgctgagatgctctctctcggatattt
239
ANTISENSE
Score: 100
Oligos:
All
Oligo Name
Type
Temp
Bases
Sequence
1234-45-1-3
Invader
76.47 °C
[31]
ctagatcgcgcgctggatagatatagcgcgT
Done
Local intranet

FIGURE 14F

Design Viewer - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://jupiter/servlet/jsp/DesignViewer.jsp?mode=Transgene&orderKey=-9163785075533198484

1234-45-1-1	Probe	64.29 °C	[16]	CGCGCCGAGGctagagatcgcgctgg
1234-45-1-2	Synthetic Target		[52]	CTATCCAGCGCGATCTCTAgCGCGCTATATCTATCCAGCGCGCGATCTAGCG

Notes:

ANTISENSE
Score: 101
Oligos:

Oligo Name	Type	Temp	# Bases	Sequence
1234-45-2-3	Invader	76.70 °C	[30]	cgcgctttagagatgcgctcgtataggctT
1234-45-2-1	Probe	61.17 °C	[15]	CGCGCCGAGGCcgcgctggatatag
1234-45-2-2	Synthetic Target		[50]	AGAGCTATATCCAGCGCGgAGCCTATACGAGCGCATCTCTAAAGCGCGAC

Notes:

Job Submit | Printable Page / Job Submit | Cancel

Done | Local intranet

Figure 15

Automated primer selection for multiplex PCR using *Invader™ Creator Primer Designer*

Multiplex PCR commonly requires extensive optimization to avoid biased amplification of select amplicons and the amplification of spurious products resulting from the formation of primer-dimers. In order to avoid these problems, we have designed *Invader™ Creator Primer Designer* software for the automated selection of multiplex primers. Beginning with a set of user defined sequences and corresponding SNP locations, *Invader ™ Creator Primer Designer* defines an "Invader™ footprint" (the minimal amplicon required for Invader™ detection) for each sequence. Primers are designed outward from the "Invader™ footprint" and evaluated against several criteria, including the potential for primer-dimer formation with previously designed primers in the current multiplexing set. *Invader ™ Creator Primer Designer* continues through multiple iterations of the same set of sequences until primers against all sequences in the current multiplexing set can be designed.

A.

29043,FMO1,aagttagaagaaccaagactatcttgtcaggggtgtattttgagagtggcagacttttcagtgcct
ttccattcatgacacttcttgaatctctggcagaaccagccagccgtgttcacagtgtcaaatgaagggatgtcttt
gattgcttccaggtgttcctcagcaccaccggagggggatgggtgatcagccgaatctttgactcgggctacccatg
ggacatggtgttcatgacacgctttcagaacatgttgagaaattccctcccaac[ct]ccaattgtgacttggttga
tggagcgaaagataaacaactggctcaatcatgcaaattacggcttaataccagaagacaggtaaatataatgtgac
tgccaagggctttaggaagaaggagcctctgcctgtccagcagcctatacaagccaggcagtaccacagcaacatg
gctgaatgtgtgggaacacttgatacaaatttgcttgataataacagctaactgttcttaagtactcagaaagtgaa
attatgtatttc

B.

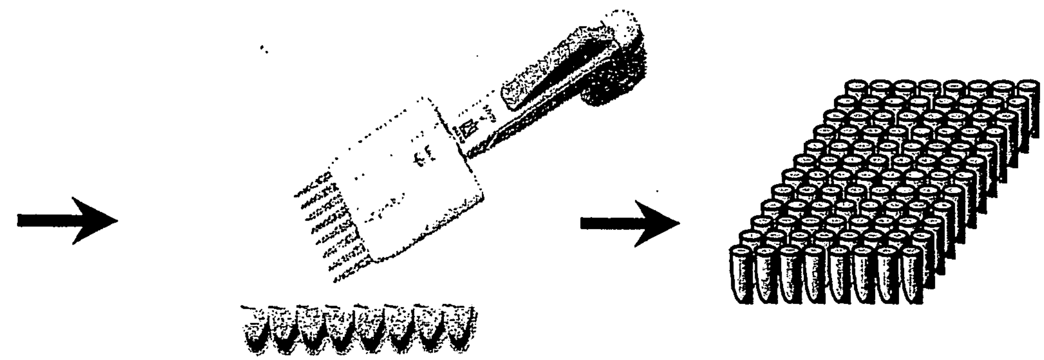
29043,FMO1,aagttagaagaaccaagactatcttgtcaggggtgtattttgagagtggcagacttttcagtgcct
ttccattcatgacacttcttgaatctctggcagaaccagccagccgtgttcacagtgtcaaatgaagggatgtcttt
gattgcttccaggtgttcctcagcaccaccggagggggatgggtgatcagccgaatctttgact**cgggctacccatg**
ggacatggtgttCATGACACGCTTTCAGAACATGTTGAGAAATTCCCTCCCAAC[ct]CCAATTGTGACTTGGTTGA
TGGAGCGAAAGATAAACAACTGGc**tcaatcatgcaaattacggcttaatacca**gaagacaggtaaatataatgtgac
tgccaagggctttaggaagaaggagcctctgcctgtccagcagcctatacaagccaggcagtaccacagcaacatg
gctgaatgtgtgggaacacttgatacaaatttgcttgataataacagctaactgttcttaagtactcagaaagtgaa
attatgtatttc

f,cgggctacccatgggaca,59.38 **r,tctggtattaagccgtaatttgcatgattga,60**

Figure 2. Creation of 101 primer sets from sequences available for analysis on the Invader™ *M*edically *A*ssociated *P*anel using Invader ™ Creator Primer Designer
(A) Sample input file of a single entry. Information includes TWT SNP#, short name identifier, and sequence with the SNP location indicated in brackets. (B) Sample output file of a the same entry. Information includes the sequence of the "Invader footprint" (capital letters flanking SNP site), forward and reverse primer sequences (bold), and their corresponding Tm's.

Figure 16

Basic workflow for highly multiplexed PCR using the Invader™ *<u>M</u>edically <u>A</u>ssociated <u>P</u>anel* (MAP)



Multiplex PCR
(10 ng / template)

Dilute and add directly to Invader™ *Medically Associated Panel*

Incubate at 63°C and read

Analysis of 101-plex PCR using the Invader™ *Medically Associated Panel*

Using primers designed by the *Invader™ Creator Primer Designer* software, highly multiplexed PCR was carried out *without* prior testing of individual "primer sets" in uniplex PCR. Of the 101 possible amplicons, 94 (~93%) were detected by the Invader™ assay and made the correct call corresponding to genomic typing of the same sample.

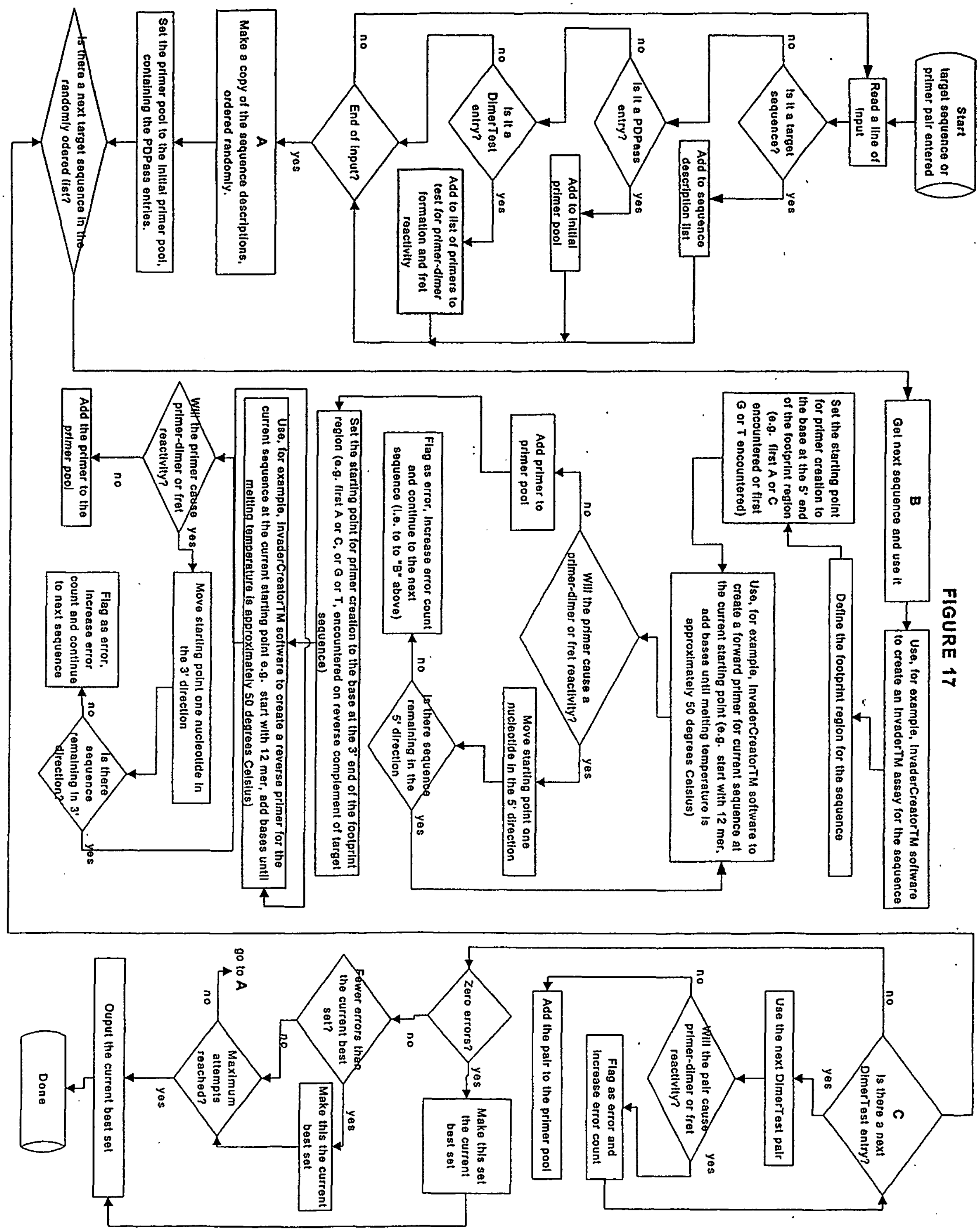

Start
target sequence or primer pair entered
Read a line of input
Is it a target sequence?
yes
no
Add to sequence description list
Is it a PDPass entry?
yes
no
Add to initial primer pool
Is it a DimerTest entry?
yes
no
Add to list of primers to test for primer-dimer formation and fret reactivity
End of input?
yes
no
A
Make a copy of the sequence descriptions, ordered randomly.
Set the primer pool to the initial primer pool, containing the PDPass entries.
Is there a next target sequence in the randomly odered list?
B
Get next sequence and use it
Use, for example, InvaderCreatorTM software to create an InvaderTM assay for the sequence
Define the footprint region for the sequence
Set the starting point for primer creation to the base at the 5' end of the footprint region (e.g. first A or C encountered or first G or T encountered)
Use, for example, InvaderCreatorTM software to create a forward primer for current sequence at the current starting point (e.g. start with 12 mer, add bases until melting temperature is approximately 50 degrees Celsius)
Will the primer cause a primer-dimer or fret reactivity?
no
yes
Add primer to primer pool
Move starting point one nucleotide in the 5' direction
Is there sequence remaining in the 5' direction
yes
no
Flag as error, increase error count and continue to the next sequence (i.e. to "B" above)
Set the starting point for primer creation to the base at the 3' end of the footprint region (e.g. first A or C, or G or T, encountered on reverse complement of target sequence)
Use, for example, InvaderCreatorTM software to create a reverse primer for the current sequence at the current starting point e.g. start with 12 mer, add bases until melting temperature is approximately 50 degrees Celsius)
Will the primer cause primer-dimer or fret reactivity?
no
yes
Add the primer to the primer pool
Move starting point one nucleotide in the 3' direction
Is there sequence remaining in 3' direction?
yes
no
Flag as error, increase error count and continue to next sequence
C
Is there a next DimerTest entry?
yes
no
Use the next DimerTest pair
Will the pair cause primer-dimer or fret reactivity?
yes
no
Flag as error and increase error count
Add the pair to the primer pool
Zero errors?
yes
no
Make this set the current best set
Fewer errors than the current best set?
yes
no
Make this the current best set
Maximum attempts reached?
yes
no
go to A
Ouput the current best set
Done

FIGURE 17

FIGURE 18

Primer Sequence (5' > 3')	Oligo	[Na+] (mM)	ΔΓo37	ΔH	ΔΣ	CT (M)	Tm	[Probe] (M)	[Target] (M)	Tm
ctgttcttcctgaagcctc	1117-75-1	75	-17.58	-146.35	-415.2	4.0E-04	64.48	1.00E-06	4.00E-14	56.5
ttgaggttggtgccttc	1117-70-2	75	-16.42	-129.29	-364.0	4.0E-04	65.10	1.00E-06	4.00E-14	56.0
aagagtgtattgagagcct	1117-75-3	75	-16.24	-140.66	-401.2	4.0E-04	62.20	1.00E-06	4.00E-14	54.0
tcagccttaaaaagacctcc	1117-75-4	75	-17.63	-150.66	-428.9	4.0E-04	63.74	1.00E-06	4.00E-14	56.0
ctcgtcacttcctctgtcc	1117-75-5	75	-18.41	-148.54	-419.6	4.0E-04	66.10	1.00E-06	4.00E-14	58.1
gggagaagtgcggcac	1117-70-6	75	-17.72	-127.64	-354.4	4.0E-04	69.33	1.00E-06	4.00E-14	59.9
gaccttatgtgtttttcc	1117-70-7	75	-14.26	-135.93	-392.3	4.0E-04	57.91	1.00E-06	4.00E-14	49.6
caatttcctcaaaagactttcc	1117-70-8	75	-17.78	-167.59	-483.0	4.0E-04	61.15	1.00E-06	4.00E-14	54.3
aaggacttaagaattgtcac	1117-70-9	75	-15.64	-149.50	-431.6	4.0E-04	59.15	1.00E-06	4.00E-14	51.6
cctcaatccttcaccgc	1117-70-10	75	-16.93	-132.72	-373.4	4.0E-04	65.73	1.00E-06	4.00E-14	56.9
gaggtagtgtttacagccc	1117-70-11	75	-17.44	-145.86	-414.1	4.0E-04	64.23	1.00E-06	4.00E-14	56.2
tcacatctcgagtgataatctc	1117-75-12	75	-18.35	-167.82	-481.9	4.0E-04	62.34	1.00E-06	4.00E-14	55.4
tgatgggagacgagttc	1117-70-13	75	-15.84	-129.58	-366.8	4.0E-04	63.39	1.00E-06	4.00E-14	54.4
tgcacacacacacatacc	1117-75-14	75	-17.30	-139.56	-394.2	4.0E-04	65.19	1.00E-06	4.00E-14	56.8
aggtcccctccgctc	1117-70-15	75	-16.65	-114.39	-315.1	4.0E-04	69.91	1.00E-06	4.00E-14	59.4
cacagtggtgttggac	1117-70-16	75	-15.32	-124.06	-350.6	4.0E-04	63.15	1.00E-06	4.00E-14	53.8
atcctgaagagcaagtcc	1117-70-17	75	-16.43	-135.68	-384.5	4.0E-04	63.71	1.00E-06	4.00E-14	55.1
attccggtttggttctcc	1117-70-18	75	-16.89	-136.61	-386.0	4.0E-04	64.74	1.00E-06	4.00E-14	56.2
ctgaaacccaggactcc	1117-70-19	75	-16.20	-129.29	-364.7	4.0E-04	64.48	1.00E-06	4.00E-14	55.4
ggaacaatcaccttttctc	1117-75-20	75	-15.82	-144.52	-415.0	4.0E-04	60.43	1.00E-06	4.00E-14	52.5

SNP53530,
AGGCTCCCTTAGATGCCTGACATTCTGTTCTTCCTGAAGCCTCACTCCCTTCTCTCCTGGctGCAGACACGTCCCCATCAGAAGGCACCAACCTCAACGCGCCCAACAGCCTGGGTGTCAGC
Primers 01/02

SNP54874
TAAAATCATTTATTTATTTATCCATCCATCAAGAGTGTATTGAGAGCCTGACAACATACCagGCATCAAGCCCTGGAGGTCTTTTTAAGGCTGAGCCAATATAGCTATGGATAACATTCTAA
Primers 03/4

SNP67356
CCCTCGTCACTTCCTCTGTCCTGTGGGGTGGGGGTGCAGGCGCTctCTCCTTTAGCTGTGCCGCACTTCTCCCTACAGGCCAGGAGAAACA
Primers 05/6

SNP47950
CTTCTGTGGACCTTATGTGTTTTTCCTCTTTGCTGGAGTGCTCCTGGCCTTTACCCTGTTctACATTTTTTAAAGTTCCAGAAACCAAAGGAAAGTCTTTTGAGGAAATTGCTGCAGAATTC
Primers 07/8

SNP41646
TGGTTAAGGACTTAAGAATTGTCACTTGTGTGTGTATATTGTTGTTGTTGTTGCAACGGTgtCTGTGTACGCACGGTTACAGTGGATCAAATTTGGGGAGTTAGGAAGTGGCGTTGGTTTGT
Primers 09/10

SNP67320
CGAGGTAGTGTTTACAGCCCTCATGAACAGCAAAGGCGTGAGCCTCTTCGagCATCATCAACCCTGAGATTATCACTCGAGATGTGAGTAC
Primers 11/12

SNP67325
GGAGACGAGTTCAAGGTGAGTGGGTGGGGCTGGGCTGCTAGGGgaATCCAGATGGCATGTGGTATGTGTGTGTGTGCACACGCATGGGGAG
Primers 13/14

SNP47901
CAGCCACAGGTCCCCTCCGCTCAGGTGATGGACTTCCTGTTTGAGAAGTGGAAGCTCTACagGTGACCAGTGTCACCACAACCTGAGCCTGCTGCCCCCTCCCACGGGTGAGCCCCCCACCC
1117-70-15/16

SNP47932
AGAGCAAGTCCCCCAAGGAGGAGCTGCTGAAGATGTGGGGGGAGGAGCTGACCAGTGAAGacAAGTGTCTTTGAAGTCTTTGTTCTTTACCTCTCGGGAGAACCAAACCGGAATGGTCACAA
Primers 17/18

SNP47892
GGAACTGAAACCCAGGACTCCGTCTCTTGCCAGTGAAAGTTATGTTAGGAAGCAGTGAGGtgGTCTAAAGCAGTATGAAAGGCAAAGAGAAAAGGTGATTGTTCCCTCTTGAATGGCCCTTG
Primers 19/20

Figure 19

A

10 min.	53530	54874	67356	47950	41646	67320	67325	47901	47932	47892
FAM	1.1	1.2	1.0	1.0	6.4	1.0	1.0	0.9	1.0	1.0
RED	1.0	1.2	1.0	1.0	5.1	2.2	1.3	0.9	1.1	1.0
FOZm (m=10)	2.0	2.4	2.0	2.1	11.5	3.2	2.3	1.9	2.1	2.0
RED+FAM-2	0.0	0.4	0.0	0.1	9.5	1.2	0.3	-0.1	0.1	0.0
Genomic Calls					HET					

30min.	53530	54874	67356	47950	41646	67320	67325	47901	47932	47892
FAM	1.9	2.8	1.0	1.0	20.5	1.1	1.0	1.0	1.0	1.0
RED	0.9	4.2	1.0	1.1	26.3	11.4	3.0	1.0	1.4	0.9
FOZm (M=30)	2.8	7.0	2.0	2.1	46.7	12.5	4.0	1.9	2.4	2.0
RED+FAM-2	0.8	5.0	0.0	0.1	44.7	10.5	2.0	-0.1	0.4	0.0
Genomic Calls	Homo F	HET			HET	HomoR	HomoR			

60min.	53530	54874	67356	47950	41646	67320	67325	47901	47932	47892
FAM	4.6	7.3	1.1	1.2	21.8	1.2	1.0	1.0	1.0	1.2
RED	1.0	11.4	1.1	1.2	34.5	25.6	8.3	0.9	2.6	0.9
FOZm (m=60)	5.6	18.7	2.2	2.5	56.2	26.8	9.4	2.0	3.6	2.2
RED+FAM-2	3.6	16.7	0.2	0.5	54.2	24.8	7.4	0.0	1.6	0.2
Genomic Calls	Homo F	HET			HET	HomoR	HomoR		HomoR	

B

240 min.	53530	54874	67356	47950	41646	67320	67325	47901	47932	47892
Genomic (RED+FAM-2)	3.10	1.76	0.68	0.43	1.60	2.07	0.80	0.89	0.85	0.85

Figure 20

Linear Relationship between Primer Concentration and Amplification Factor

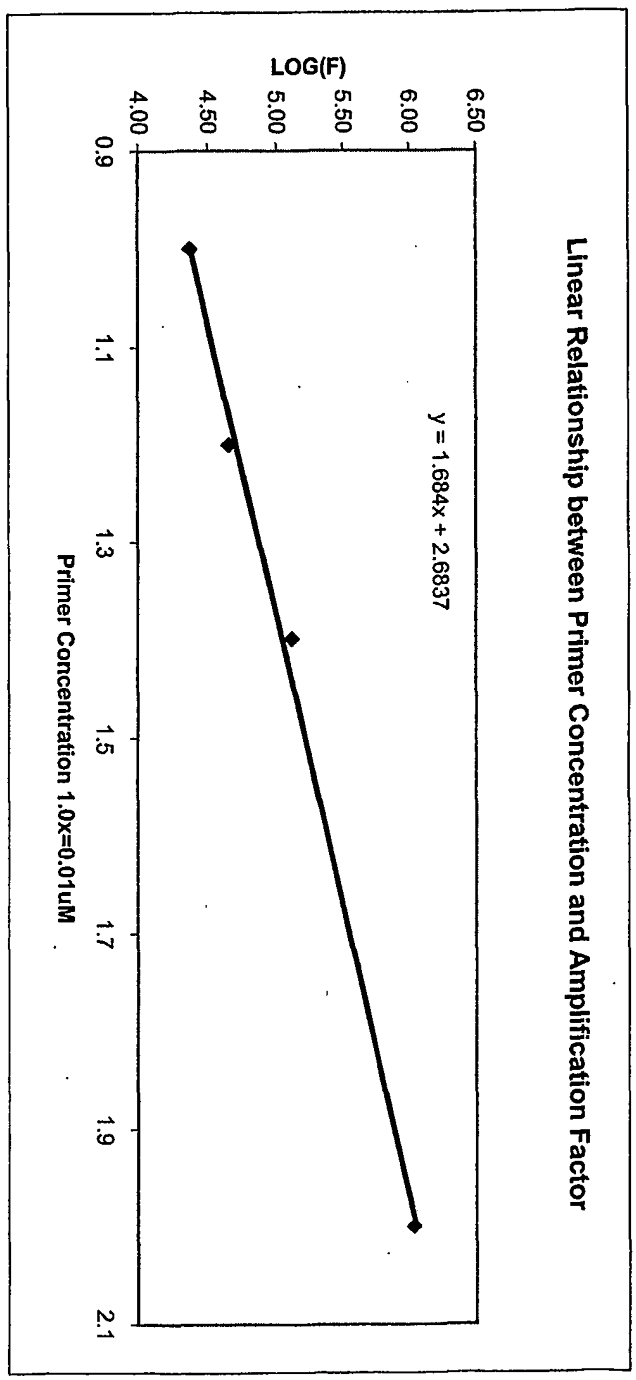


Figure 21

Amplification Factor Calculations and Apparent Primer Concentrations of Equimolar 10-plex PCR
using equation 1a'

SNP No.	m	FOZm	FOZm-2	240/m	F	Log(F)	X[n]	R[n]	Concentrations (uM) based on R[n]
53530	60	5.6	3.6	16	123552	5.09	1.4	1.8	0.018
54874	30	7	5	64	686400	5.84	1.9	1.4	0.014
67356	60	2.2	0.2	16	6864	3.84	0.7	3.8	0.038
47950	60	2.5	0.5	16	17160	4.23	0.9	2.8	0.028
41646	10	11.5	9.5	576	11737440	7.07	2.6	1.0	0.010
67320	30	12.5	10.5	64	1441440	6.16	2.1	1.3	0.013
67325	30	4	2	64	274560	5.44	1.6	1.6	0.016
47901	60	2	0	16	0	#NUM!	N/A	4.0	0.040
47932	60	3.6	1.6	16	54912	4.74	1.2	2.1	0.021
47892	60	2.2	0.2	16	6864	3.84	N/A	4.0	0.040

FOZm was chosen based on being within the dynamic range of the instrument: 0.5-12.5 for this experiment.
FOZ240 = 1.7
D=.333
given that the values of FOZ for sample 47901 and 47892 were low, an estimate of R[n] of 4 was made.
This is a good estimate, since sample 67346 is also low and the calculated R[n] value is near 4.

Figure 22

Invader FOZ values for Optimized Primer Concentration of 10-plex PCR

10 min.	53530	54874	67356	47950	41646	67320	67325	47901	47932	47892
RED	1.1	2.3	2.4	2.5	2.9	3.3	1.8	1.0	2.6	1.0
FAM	2.2	1.6	3.4	2.1	3.2	1.0	1.0	3.4	1.0	5.3
RED+FAM-2	1.3	2.0	3.8	2.6	4.1	2.3	0.7	2.4	1.6	4.3
Genomic Calls	Homo F	HET	HET	HET	HET	HomoR	HomoR	HomoF	HomoR	HomoF

Primer Concentration of 9/10 at 1.0x

FIGURE 23a

29017

CACTAGACCGCCTGTCCCCAAGGGAGCCTCAGTGGGGCGACAGGGTGCTCGGCGGACTC
CACCTCAGGCCCTCCCCACTGTTGCTGTGCATTCCTGTGCAGGTGCATCTCTTTCTTAC
TAACTGGTATTTATTAAGGGAGGTGCTCTGTAGGTCTGGAGCCTTTCCCTCATCCTTTT
TGCGAGTCCCCACCTTTTTGTTTTTTTTTTTTTCTTTGAGGCTCACTAGAGGACGCAGA
ACCTTGGGAGATTGATTTGCACAGAACTCCCCACCTCCCACTTTTACAATTTCCAGTTT
CTGATTGAAAATTTTAGGGTTTCTCCCCACTGCCCTTCCCTATCTTTCCTTCCCCTCAA
CACCATGAAGGAAAAACACACACGGCAGGGCTTTTGTAGCCCTGAAGGCAACTTTAGA
CATTTAAAATCCAGCACTTTAATCTCTTGTTCTCTGTGAATCACTATGAGAAGTGAATG
GTTTTAAAGGCTGTAATGCTATGTTGGAAATTGGTTTGTTTTGCCTTTTATTGAAAAGG
TAAGATCATGTGATTGGAAGAACACAACT[gt]TTGGCTTGGGAAGAGGACTTTGCTGC
TGAAGTGTTTTCTACCTTCTGAGTGTGTTTAAGGCAGGATTTGGAGGGAAGGACCAGCT
TAGGGAGAGTGTCTGAGCCACAGCGTCAGGATGGGGAAACCACATGGGATCCATCAAG
TTCCAGTTGAACAGGAGCAAGATCAGAACTTAGGAGGGCAGTGTCAGCTCCCTTGTTGG
CTGTCAAGGAACACCGATCTAGTAGAAACCCACTTGGTTGTGACCCAGGTAGAGGTAGA
TGCCATACATTTGAGATATGCGTCCTTAAGGAACCTGACAAGCAGACTGAAGGGATGGT
AAGTGTGACAGCCTGATAAGTTTTCTCAAAGCCCAGGATACAGAGCCAGTGTTTTCTGT
AACTGGAGACCTCAGTTAGGCCAACTTCGAATTCCAGAGCAACGTAGGAAGTCTATTCA
GCAGAAACTCGACATTGTTCA

cactagaccgcctgtccccaagggagcctcagtggggcgacagggtgctcggcggactc
cacctcaggccctccccactgttgctgtgcattcctgtgcaggtgcatctctttcttac
taactggtatttattaagggaggtgctctgtaggtctggagcctttccctcatcctttt
tgcgagtccccacctttttgtttttttttttttctttgaggctcactagaggacgcaga
accttgggagattgatttgcacagaactccccacctcccacttttacaatttccagttt
ctgattgaaaattttagggtttctccccactgcccttccctatctttccttcccctcaa
caccatgaaggaaaaacacacacggcagggcttttgtagccctgaaggcaactttaga
catttaaaatccagcactttaatctcttgttctctgtgaatcactatgagaagtgaatg
gttttaaaggctgtaatgctatgttggaaattggtttgttttgccttttattgaaaAGG
TAAGATCATGTGATTGGAAGAACACAACT[gt]TTGGCTTGGGAAGAGGACTTTGCTGC
TGAAGTgttttctaccttctgagtgtgtttaaggcaggatttggagggaaggaccagct
tagggagagtgtctgagccacagcgtcaggatggggaaaccacatgggatccatcaag
ttccagttgaacaggagcaagatcagaacttaggagggcagtgtcagctcccttgttgg
ctgtcaaggaacaccgatctagtagaaacccacttggttgtgacccaggtagaggtaga
tgccatacatttgagatatgcgtccttaaggaacctgacaagcagactgaagggatggt
aagtgtgacagcctgataagttttctcaaagcccaggatacagagccagtgttttctgt
aactggagacctcagttaggccaacttcgaattccagagcaacgtaggaagtctattca
gcagaaactcgacattgttca

FIGURE 23b

29043
ATACCAAAAGTAATTGTAGTACTGAATTTTGCTGTCATTTAAGCCAATGGTTTGCACTG
AAACTCTGTAGACAACTCTGATACTGCCATTCCCTGTTCTTACTGCCTACAATGATAGT
GAGCACACCAAGTAGCAATCACCTGTTCATTGTTTTCTTACATAGACTTTAGGTCCCTA
TGGTTTACTAAAGGCTGGCAGATAATAAGTATTCAATAATATGTCTTAAGGCATTTTAA
TACTCTAGATGCTCTGAATCCTAATCTCAAAAGGATTAACTTTAAAATAGAAGTTAGAA
GAACCAAGACTATCTTGTCAGGGGTGTATTTTGAGAGTGGCAGACTTTTCAGTGCCTTT
CCATTCATGACACTTCTTGAATCTCTGGCAGAACCAGCCAGCCGTGTTCACAGTGTCAA
ATGAAGGGATGTCTTTGATTGCTTCCAGGTGTTCCTCAGCACCACCGGAGGGGGATGGG
TGATCAGCCGAATCTTTGACTCGGGCTACCCATGGGACATGGTGTTCATGACACGCTTT
CAGAACATGTTGAGAAATTCCCTCCCAAC[ct]CCAATTGTGACTTGGTTGATGGAGCG
AAAGATAAACAACTGGCTCAATCATGCAAATTACGGCTTAATACCAGAAGACAGGTAAA
TATAATGTGACTGCCAAGGGCTTTTAGGAAGAAGGAGCCTCTGCCTGTCCAGCAGCCTA
TACAAGCCAGGCAGTACCACAGCAACATGGCTGAATGTGTGGGAACACTTGATACAAAT
TTGCTTGATAATAACAGCTAACTGTTCTTAAGTACTCAGAAAGTGAAATTATGTATTTC
ACCTTGTCAGCAACACTTTACGTATTATTATAATAATCCTTTTATTATGGAGAAACTGA
AACAGCAAAATTCAGCCATTTACCCAAGCTCACTGAGTAGTAAGTGAACTCTGTGACCT
TGGCAAGTTACTTGATCCTCAGCTGTAGCAACCAAAAGAGAATGATTTGTCTATGACTT
TGTTGATAAAAGAAACACACT

ataccaaaagtaattgtagtactgaattttgctgtcatttaagccaatggtttgcactg
aaactctgtagacaactctgatactgccattccctgttcttactgcctacaatgatagt
gagcacaccaagtagcaatcacctgttcattgttttcttacatagactttaggtcccta
tggtttactaaaggctggcagataataagtattcaataatatgtcttaaggcattttaa
tactctagatgctctgaatcctaatctcaaaaggattaactttaaaatagaagttagaa
gaaccaagactatcttgtcaggggtgtattttgagagtggcagacttttcagtgccttt
ccattcatgacacttcttgaatctctggcagaaccagccagccgtgttcacagtgtcaa
atgaagggatgtctttgattgcttccaggtgttcctcagcaccaccggagggggatggg
tgatcagccgaatctttgactcgggctacccatgggacatggtgttCATGACACGCTTT
CAGAACATGTTGAGAAATTCCCTCCCAAC[ct]CCAATTGTGACTTGGTTGATGGAGCG
AAAGATAAACAACTGGctcaatcatgcaaattacggcttaataccagaagacaggtaaa
tataatgtgactgccaagggcttttaggaagaaggagcctctgcctgtccagcagccta
tacaagccaggcagtaccacagcaacatggctgaatgtgtgggaacacttgatacaaat
ttgcttgataataacagctaactgttcttaagtactcagaaagtgaaattatgtatttc
accttgtcagcaacactttacgtattattataataatccttttattatggagaaactga
aacagcaaaattcagccatttacccaagctcactgagtagtaagtgaactctgtgacct
tggcaagttacttgatcctcagctgtagcaaccaaaagagaatgatttgtctatgactt
tgttgataaaagaaacacact

FIGURE 23c

34573
CAGCTGTGGGGTCAGGAAGGGCTTGAAGTATGGGACACTAGCCTGCCCCACCTCCACTC
TGCAGCACCCACAGGACCACCCTCATGCCCCTGGCAACAGCATGCAGGGCAGCTGCAGG
ATCCAGGTGGGACCCAGATACTATATGAAGGAGCCACCTTACCTGCTTTTTGCAAAGCT
ACTGGGATGGCATAGGCAGGTCCAATGCCCATGATGTCAGGTGGGACCCCAACCACTGC
ATAAGACCTCAGGACCCCAAGGATGGGAAGGCCCAACTCTTCTGCCTTGGACCTCCGGG
CCAGCAGGATGGCAGCTGCCCCATCACTCACCTGGCTAGAGTTTCCTAGGGGCAAACTG
TTGGGGTAAGAAGGCATCGGGGTGGGGATGAGGAGATCCCAGCCCTCCCACTTCTACTT
TGCAGAGGGGCCTGGTCTATTCCANGTTCCCAGAGTACAGCACCCAGCATGGCCATGGC
CTGCTTTCTCATACCCCTACCCCGGACCAGTNTCACCAGCTGTGGTAGAACCATCTTTC
TTGAAGGCAGGCTTCAGTTTGGCCAGGCC[ct]TCCATGGTGGTGCTGGGGCGGATACC
CTCATCCTGGGTCACAGTGATGCTCCTCTTGGTGCCCTTGTCATCATGGACCGTGGTGG
TCACAGGCACAATCTCAGCTTGGAAACAGCCCTTGCTCTGGGCTCTTGCTGCCCTGCCA
GCACCATGGACAGCCAGCTTCAGACTCCCTTGGGGTTCCCTTCCTTCCCTGCCCCCAAC
CCCTATCCATTTGGGTAGACACAAGCTCAGGCTGCTAAATTCAGGGACATGCTCGACTT
TGGGGGAGCTCTGAGGGCATGGCTAAGGCCTTACAGGGCCTTCTTCACCATCAGCCCCA
GACCTCCAGATCGTGGCCAATCCCAACCTCAAAGGGGGGAAAGGGTGTTTGGAAGTGGT
GCCTCCACTTAGAGCCCTTTGTCCAAGAGGGATTAAGCCTGCTTGATTCTCTCTGCTAA
ACTGAGGATGGAACCCCAGAA

cagctgtggggtcaggaagggcttgaagtatgggacactagcctgccccacctccactc
tgcagcacccacaggaccaccctcatgcccctggcaacagcatgcagggcagctgcagg
atccaggtgggacccagatactatatgaaggagccaccttacctgctttttgcaaagct
actgggatggcataggcaggtccaatgcccatgatgtcaggtgggaccccaaccactgc
ataagacctcaggaccccaaggatgggaaggcccaactcttctgccttggacctccggg
ccagcaggatggcagctgccccatcactcacctggctagagtttcctaggggcaaactg
ttggggtaagaaggcatcggggtggggatgaggagatcccagccctcccacttctactt
tgcagaggggcctggtctattccangttcccagagtacagcacccagcatggccatggc
ctgctttctcatacccctaccccggaccagtntcaccagctgtggtagaaccatctttc
ttgaAGGCAGGCTTCAGTTTGGCCAGGCC[ct]TCCATGGTGGTGCTGGGGCGGATACc
ctcatcctgggtcacagtgatgctcctcttggtgcccttgtcatcatggaccgtggtgg
tcacaggcacaatctcagcttggaaacagcccttgctctgggctcttgctgccctgcca
gcaccatggacagccagcttcagactcccttggggttcccttccttccctgcccccaac
ccctatccatttgggtagacacaagctcaggctgctaaattcagggacatgctcgactt
tgggggagctctgagggcatggctaaggccttacagggccttcttcaccatcagcccca
gacctccagatcgtggccaatcccaacctcaaaggggggaaagggtgtttggaagtggt
gcctccacttagagccctttgtccaagagggattaagcctgcttgattctctctgctaa
actgaggatggaaccccagaa

FIGURE 23d

37319
GAGGATCAAAGCACCTGGTACAATGCCTGGCCAGAAAGTTGAATAATCGAATATAGCTA
ACGTCACTATTGCAGGCTGGCTATGTGCCTGGCGGTGTTCTTAGCCATTTACAAGTATG
AACTCATTTAATCCTCATAAGATCCTGTATGAGGTGAGTAAGCTGTTAATTCCCTTCCT
TGCCCATACTCTGTGACTCCAACCCACCACAGTTGAATTTCTCCTTATGAATTATAAAT
CAGAAAACGGCCCCAAATTCTGTCATGTCTAAGTGGGAAAATGGAAGAAGGCATTGATT
TCTCCCCTACTCAAGCAGAAGAGAATTAACCTCAGTCCCTGCTTTGCCCATATTCCTTC
CCCAGGGCCCCAGGAAGAAGACATGGAAAAACAATATTTCCACCAAAGTTTATTTCTCT
GAAACAATCACCAGTTGCTGTCCTCTATGGCACACTGAGAGCCCCAGGAGGGTCTTTAA
CTCCCTTCCTCAGATTATATTCATCCCAGAAATATAGCCTTGGACAATAATTTGGTTAC
AGCATAGTCCCAGGAATGAGGTCCCCCAA[ga]TTGCTAAGTTTTACATAGGGGAGACT
GGGAAATTCAAAGAATTGGATGGAGAAACCATAGGATCCAAGATAATGTCAGGGGGTTG
AAGATGTTGGAGAGGCATGGTAGCATCATTGAGTTTGAATCTCCTTCTCACTTGGAGTG
GAAGTTGTAGGATTCTGCCTCTAGGAAATGTGCCATCCTACAGAATAAATAAAAGGGAG
ATAATGAGGCTTCAACCCAACTTGCCCCCATCGTTTGTCACTGTAACCATCCCATGCCT
TAATACAGTGATACTGAAAACTCCAGGGCACCAACAACTAATACAAAGGAAGCACCTTC
AGCCTCCTCTCCACAGACATCCCACTTGGTAGAAGAGGAGGATGCTCCTTCCTGCTCTT
AATCCTAGCAATGGCAGCTTAAATCATGCCCTTGCCTAGATCCTCATGGAAGCTCACCC
ATATAATAATCAAGATTAGTT

gaggatcaaagcacctggtacaatgcctggccagaaagttgaataatcgaatatagcta
acgtcactattgcaggctggctatgtgcctggcggtgttcttagccatttacaagtatg
aactcatttaatcctcataagatcctgtatgaggtgagtaagctgttaattcccttcct
tgcccatactctgtgactccaacccaccacagttgaatttctccttatgaattataaat
cagaaaacggccccaaattctgtcatgtctaagtgggaaaatggaagaaggcattgatt
tctcccctactcaagcagaagagaattaacctcagtccctgctttgcccatattccttc
cccagggccccaggaagaagacatggaaaaacaatatttccaccaaagtttatttctct
gaaacaatcaccagttgctgtcctctatggcacactgagagccccaggagggtctttaa
ctcccttcctcagattatattcatcccagaaatatagccTTGGACAATAATTTGGTTAC
AGCATAGTCCCAGGAATGAGGTCCCCCAA[ga]TTGCTAAGTTTTACATAGGGGAGACT
GGGAAATTCAAAGAATTGGATGGagaaaccataggatccaagataatgtcagggggttg
aagatgttggagaggcatggtagcatcattgagtttgaatctccttctcacttggagtg
gaagttgtaggattctgcctctaggaaatgtgccatcctacagaataaataaaagggag
ataatgaggcttcaacccaacttgcccccatcgtttgtcactgtaaccatcccatgcct
taatacagtgatactgaaaactccagggcaccaacaactaatacaaaggaagcaccttc
agcctcctctccacagacatcccacttggtagaagaggaggatgctccttcctgctctt
aatcctagcaatggcagcttaaatcatgcccttgcctagatcctcatggaagctcaccc
atataataatcaagattagtt

FIGURE 23e

37474
AGGTGCACTTTTTCCAGGACCTCCTGCACAGGTGTGATATTTAGCCTGGAAGCAATGTG
TACATGGAATGCCCTACAGGCACAGGAGGCATCCCTGGAGACTGAATGGTGTCTGGGAA
GAGTAGGGCCACAGAGCTGAGCCCCTATGGACTGCAGCAGAGGGCCTGGCTCCAATCCT
AGCCTACCATATCCCAGTCCCATGATCGTGAGTAGTCCCATGGGATCAAGTGCTCTCAT
TCATAAAAGAAGGGAGGTAACAGCTGCCCCACTCACGCCCCAGGATCATCCGGCAGTCA
AAGGGGATTCAGGTGCTTCCTGGAAGACAGAGTCACAGGGGACCCTCCTTTTCCCAGCC
ACCCATATCAGTCCACCTTTTGGGTTTTGACCTTTACTATGTGGTTTTCTAGACTTCTA
TTGACAAATCCTGCTTTATGGACAGGGATGCTTTTCATTTAGATTGGGGGCCACTCCCC
AACATCTCATTTATTTTTCACAGCTCTGGTCCCATGGAGTCTTGTTTGAGTGCAAGTGA
ACTGAATTTCCCAATTCCTCAAAAAGAGC[ca]ATAGTAATAAAAACCATAATAGTGAC
ACTTACATATGGATAGTGCTTTGTAGTTTAGAAAATGCTTTCACCAACTGATTGCCATG
ACAGCCCTGAGAAGTAACCTACTCTACAGATGAGGAGCCTAGAGAGAGAAAGTGACTTT
CCTGGGCACATAGGCCCATGAGGTTCTGGTGCCAGCATAATAGACTAGTCAAATTTCCA
GACTCTGGAGTCAGACTGCCTGAGTTCAAACCATGGGTCCTCTTGGTCAGGTTTTATAA
CCACTCTAAAACTCTGTTTGCCCATCTGTAAAGTGAGCACAATTACAGAATCTACCTAA
TAGGGCTGTCTGTATGTCAATGGGCTTGGCCTGTGCCTGAGGAAATGCTANCCCCATGA
TCCTGCAGCCATGGTTAGGAAGGACATGGCAGGGAATGGGACCTTTCACAGACCGGGCT
GTGGCCAGCAGCCAGGGCCGA

aggtgcactttttccaggacctcctgcacaggtgtgatatttagcctggaagcaatgtg
tacatggaatgccctacaggcacaggaggcatccctggagactgaatggtgtctgggaa
gagtagggccacagagctgagcccctatggactgcagcagagggcctggctccaatcct
agcctaccatatcccagtcccatgatcgtgagtagtcccatgggatcaagtgctctcat
tcataaaagaagggaggtaacagctgccccactcacgccccaggatcatccggcagtca
aaggggattcaggtgcttcctggaagacagagtcacaggggaccctccttttcccagcc
acccatatcagtccaccttttgggttttgacctttactatgtggttttctagacttcta
ttgacaaatcctgctttatggacagggatgcttttcatttagattgggggccactcccc
aacatctcatttatttttcacagctctgGTCCCATGGAGTCTTGTTTGAGTGCAAGTGA
ACTGAATTTCCCAATTCCTCAAAAAGAGC[ca]ATAGTAATAAAAACCATAATAGTGAC
ACTTACATATGGATAGTGCTTTGTAGTTTAGAAAatgctttcaccaactgattgccatg
acagccctgagaagtaacctactctacagatgaggagcctagagagagaaagtgacttt
cctgggcacataggcccatgaggttctggtgccagcataatagactagtcaaatttcca
gactctggagtcagactgcctgagttcaaaccatgggtcctcttggtcaggttttataa
ccactctaaaactctgtttgcccatctgtaaagtgagcacaattacagaatctacctaa
tagggctgtctgtatgtcaatgggcttggcctgtgcctgaggaaatgctancccccatga
tcctgcagccatggttaggaaggacatggcagggaatgggacctttcacagaccgggct
gtggccagcagccagggccga

FIGURE 23f

37810
TCATACAACTCCTTGCAGTTCATGTAAGGACTCGGATTTTACCTGGAGTGGAAAAAGAA
GCACTGAAAGATTTGAGCAGGGGAGTAACCTGATAGCGTTTATGTTTAGTCCTGCCACT
TCGACAGATAAACGCACCAATGGGCTTGATGAGATTTAGGCCAACCCATAACCGCCCCT
CAACTTCTTTCCTTTCAATTTCAAAACTCCTCTATGGCTTCCTCCATCTGTTCTTCCTT
CTGAGAAGTGCTCTCTCTGCCCCTTTACAGAACTAACCACTTCGGCAACTCCTTGGACA
CTTTCCTTCTTGTTAATAATTTGCTTTCTCCGCCCCTCAAAAGCTTGCTGTTTCTGTAA
ATCATTACCTGTAAGAGGAACCGCTGGGAGTCCTGTAAACTTTAGCCCAGAGCTTGGCT
CCTCCTCCAGAATGTCTCCACCAATCAAGGAAAGTGTTTTGGGCCAGTCTTGCTCCTCC
GGATTGTCAGACTGCTCCTCCCTCTTCTTTAGACTGCCACGAGGAAAAAGCAGATGTGA
GAACTCAAGGTTCAGGGCTGCTCTTCTAA [cg] AAACAAGTCTGCCATAATCTCCATCT
GTGTTGGAATCTGTTAACTAGTGAGTACCTCATCTCCCCTCCTGTGTAAGATTTCCTGA
ACTGGCACATCTGTTTTTTGAGCAAAGATAACAAACAGATGAACAAAACCAACAATCAA
AAATGCTGTCATTAAAGTCTTGGGCAGCCAAAGTTTCTCTCAGAATTTCTCAGTTGTGT
GATACTATCTATTAAGTGATGAGGAGTATGCACACACAAAAGGCTATAAATGTAGCAGC
TGAGTTTTCATGTTGAGCCTTTTGGTGCTATTTGATTTTTTGAAAAACTATGTACATGT
ATTAAGTTGATAAATTTTTTTTTAATTTAATTGAACCAGATGCGGTGGCTCAAGCCT
GTAATCCCACCACTTTAGGAGGCTATGGTGGGCAGATGCAGATCACTTGAGGCCAGGAG
TTCGAGACCAGCTTGGCCAAC

tcatacaactccttgcagttcatgtaaggactcggattttacctggagtggaaaaagaa
gcactgaaagatttgagcaggggagtaacctgatagcgtttatgtttagtcctgccact
tcgacagataaacgcaccaatgggcttgatgagatttaggccaacccataaccgcccct
caacttctttcctttcaatttcaaaactcctctatggcttcctccatctgttcttcctt
ctgagaagtgctctctctgcccctttacagaactaaccacttcggcaactccttggaca
ctttccttcttgttaataatttgctttctccgcccctcaaaagcttgctgtttctgtaa
atcattacctgtaagaggaaccgctgggagtcctgtaaactttagcccagagcttggct
cctcctccagaatgtctccaccaatcaaggaaagtgttttgggccagtcttgctcctcc
ggattgtcagactgctcctccctcttctttagactgccacgaggaaaaagcAGATGTGA
GAACTCAAGGTTCAGGGCTGCTCTTCTAA [cg] AAACAAGTCTGCCATAATCTCCATCT
GTGTTGGAATCtgttaactagtgagtacctcatctcccctcctgtgtaagatttcctga
actggcacatctgttttttgagcaaagataacaaacagatgaacaaaaccaacaatcaa
aaatgctgtcattaaagtcttgggcagccaaagtttctctcagaatttctcagttgtgt
gatactatctattaagtgatgaggagtatgcacacacaaaaggctataaatgtagcagc
tgagttttcatgttgagccttttggtgctatttgattttttgaaaaactatgtacatgt
attaagttgataaatttttttttaatttaattgaaccagatgcggtggctcaagcct
gtaatcccaccactttaggaggctatggtgggcagatgcagatcacttgaggccaggag
ttcgagaccagcttggccaac

FIGURE 23g

37879
TATGTGTTGAATGAAAGGCTGGGTCATATGTGACCCTTGTGAGCAGCTGTTTCCGTGGA
CTGCTCCTGGGTCCCCTCCTCCACCCGCCCTGCCTCTCCCATTTCATCCTAGGAGGTGC
CTGTGGCCGGGCGCAGTAGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGG
GCGGACCACCTGAGGTCAGGAATTTGAGACTAGCCGGCCCAACATGGCGAAACCCCATC
TCTACTAAACATACAAAAAATTAGCCAGGCGTCGTGGCGGGCGCCTGTAATCCCAGCTA
CTCAGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGCTTGCAGTGGGCC
GAGATTGCGCCACTGCACTCTAGCCTGGGGGACAACAGCGAAACTCCGTCTCAAAAATA
TATATATATATTAATTAAATAAAAAAACGAGGTGCCTTCTCCTGACTCCCTGATCCCCG
CGCTCTCCAGCTCTGCCCTCGCGATCGCTGGAGCCCCCTGAGGAACTCACGCAGACGCG
GCTGCACCGCCTCATCAATCCCAACTTCT[at]CGGCTATCAGGACGCCCCCTGGAAGA
TCTTCCTGCGCAAAGAGGTGCCGAGCACAGCCGTAGCCAGGGGAGGGGCTGAAGCGGGG
CAGGGGAGGGGCTGAAGCGAGCAGAGGAGGGTCTAGGACTTGGGGAGGGAGCCCAGGAG
GACAGAAAAAGGCCGGGCTGAAACCAGGGGTGGGGTTACAGCCGGGGCGGAACTGCATT
TAGGGGGCGGGGCCGGGTGTGAAGCAAGGCCAGGGGGCAGTCGGACAGTACCCACTGAA
GCCCCGCCCCTGCAGGTGTTTTACCCCAAGGACAGCTACAGCCATCCTGTGCAGCTTGA
CCTCCTGTTCCGGCAGGTGAGGTCCTGTCTCCCCTTTCTGCCTCAGTGAACTCAGCAGG
GCTGTGTGGACGCAAAGATGAGCTAGCTGCAAAGCCTGCCTCTGCATGTTGGGATTTGG
GGTCCTTGACAGGGGTGAGGA

tatgtgttgaatgaaaggctgggtcatatgtgacccttgtgagcagctgtttccgtgga
ctgctcctgggtcccctcctccacccgccctgcctctcccatttcatcctaggaggtgc
ctgtggccgggcgcagtagctcatgcctgtaatcccagcactttgggaggccgaggcgg
gcggaccacctgaggtcaggaatttgagactagccggcccaacatggcgaaaccccatc
tctactaaacatacaaaaaattagccaggcgtcgtggcgggcgcctgtaatcccagcta
ctcaggaggctgaggcaggagaatcgcttgaacccaggaggcggagcttgcagtgggcc
gagattgcgccactgcactctagcctgggggacaacagcgaaactccgtctcaaaaata
tatatatatattaattaaataaaaaaacgaggtgccttctcctgactccctgatccccg
cgctctccagctctgccctcgcgatcgctggagcccccctgaggaactcacgcagacgcg
gctGCACCGCCTCATCAATCCCAACTTCT[at]CGGCTATCAGGACGCCCCCTGGAAGA
tcttcctgcgcaaagaggtgccgagcacagccgtagccaggggaggggctgaagcgggg
caggggaggggctgaagcgagcagaggagggtctaggacttggggagggagcccaggag
gacagaaaaaggccgggctgaaaccaggggtggggttacagccggggcggaactgcatt
tagggggcggggccgggtgtgaagcaaggccagggggcagtcggacagtacccactgaa
gccccgcccctgcaggtgttttaccccaaggacagctacagccatcctgtgcagcttga
cctcctgttccggcaggtgaggtcctgtctcccctttctgcctcagtgaactcagcagg
gctgtgtggacgcaaagatgagctagctgcaaagcctgcctctgcatgttgggatttgg
ggtccttgacaggggtgagga

FIGURE 23h

41064
CTATATGCTTGAAAGAATTTATAATTAAAATTTTTTTAAAAAAAGAGCATGAAGACTT
GCACAGCAAGATATCAGAAAGCTAAATGGAAATTTTCTTCTTAGCTATGTGAAAGACAC
AGGCAGAGCACCAGATGGTTCAGTAGCCTGAGTTCTAGAAATAATCTCAACATGGTAAG
AGGGTCTGTAAGCTAGCCTACACCTATGCGAAACAGGGTTTTATGCATGGGACACTATT
CCAGTAGAAAATGCAGGATTTGAGTAGACTTCTAGAGTTGGTTTTAAAATGATTTAATG
TAAGGCATCAAATCTAGACAATCAGTAAGAGAGTAACCCATACAGGCTATATTTTCACA
TGTTCTATAAAGTATAGTTTGGTGTCTACAGCCTGCAAACCACAGCCAGGCCCCAAATC
TTTCAAGTTGGCCCCTGACTCTTTCCTGCTGTCTCCATATGACCGAGTATGCACTGAAC
TATCAGCGTTTCCAGGTTCCTCTCCAGGCACCGCAGAGTGGTGGCGCTCTCACAAAGGC
ATGACAGGAAGACAGGGTGTGAGGTTGGA[tc]GGAGAGAGGCTGTAGCTGAGGAAAAG
CACAGCCCATGGCATTTTACTGTAATGCCTGAACAAATGCACTTAATGAATATGTGGCA
AATGTAGGCTCAGAAGTATCATTTCTTTCCTGTAAATGTAAATGCTCTCCCTCTGAAGT
TCCTGTGGGAATGGCTTCTGGATTCTGGGGGTGAGTGTGGGGCCACCCTCCACGAGGCC
TCTGCCTACCTGAAAGCATCATTCCATAGACCCTCCCATTGTTCACACACAGTGGACCT
AACTCTCCACTTTCACTTTTTCTTCTGTAATAGTTTATAACAGTCAATAGAACTCCCAC
ATTAGCTTTTAGGGTCATCACAGAATACAAAATGTTGAAGATACATATTTTATCTTTTC
TATCTTTCTCCTTAGTATCCAGGTACACTAACTCTGATATTCTAACAGAAATTATACAG
ACACCATGATCACCATCTTGA

ctatatgcttgaaagaatttataattaaaatttttttaaaaaaagagcatgaagactt
gcacagcaagatatcagaaagctaaatggaaattttcttcttagctatgtgaaagacac
aggcagagcaccagatggttcagtagcctgagttctagaaataatctcaacatggtaag
agggtctgtaagctagcctacacctatgcgaaacagggttttatgcatgggacactatt
ccagtagaaaatgcaggatttgagtagacttctagagttggttttaaaatgatttaatg
taaggcatcaaatctagacaatcagtaagagagtaacccatacaggctatattttcaca
tgttctataaagtatagtttggtgtctacagcctgcaaaccacagccaggccccaaatc
tttcaagttggcccctgactctttcctgctgtctccatatgaccgagtatgcactgaac
tatcagcgtttccaggttcctctccaggcaccgcagagtggtggcgctctcacaaaGGC
ATGACAGGAAGACAGGGTGTGAGGTTGGA[tc]GGAGAGAGGCTGTAGCTGAGGAAAAG
CACAGCccatggcattttactgtaatgcctgaacaaatgcacttaatgaatatgtggca
aatgtaggctcagaagtatcatttctttcctgtaaatgtaaatgctctccctctgaagt
tcctgtgggaatggcttctggattctgggggtgagtgtggggccaccctccacgaggcc
tctgcctacctgaaagcatcattccatagaccctcccattgttcacacacagtggacct
aactctccactttcactttttcttctgtaatagtttataacagtcaatagaactcccac
attagcttttagggtcatcacagaatacaaaatgttgaagatacatattttatcttttc
tatctttctccttagtatccaggtacactaactctgatattctaacagaaattatacag
acaccatgatcaccatcttga

FIGURE 23i

41164

GGTTCACTCACCCCTCCTCCCACCTCGGCAGCCCTGGGATGTCGCTGCTGACTCAGGAG
GAACCCGAGGTGCCGTAGCGGCTGCTCCAATATTGCAGAAGAGGTTCCTCAGGCAGCTC
TGCCCACAGCCCCAAGTCACGAATTCCGTGACTCCAGCTCCATCCCAGGCCCCAGGGTA
CCTGGCCCAGGGTTGTGCTGCCGCAGACTTGGCCTGTACCATCCAGGCGGCGGTGGGGA
GCTGGGGTTGGAAAGGCTTCTTGGAGTGGACTCCTGGGTCTGTCTGGGAGACGGGGAGG
AAGGGACACTCTGAACATCACCAGGGGCTGCTGGGGGGCCCTGGCCACCCCCAGAGTCA
GAACAGGCAGGTGGGGCAGGATCTCAGGTCATCCTATGCTACACTCAGCCATTGCGTGG
CCCCTCTCCTCCCTGTGCCTGGCCTTTTGGCCAGCCCTGGGGCCACCGAGAGGATGCAG
CACCGAACCCTCCAGGAGCCCCCAGTGCTGCCGTCTGTGGGACAGGGACAATCCCATCC
CCACTGCTACTGTCTGTGCTGTGCTGGGC[ga]CAGAGCTGGACACCTCCAAGGCCCAG
CGCCCGTAGTGGCTCTCATCATGGACAATTCACAGGCAGATGGTGGCCAGCTCTGTGGC
CTGCAGGGACTGGGAGCGGCGCCAGACCATCTAGGCCCCAACCTATCTGCATTATCCTG
GAAGACTTCCTGGAGGAGGCTTCTAAGCTGAGGCCCAAGGACCATGTCAGGTCTAGGAC
TAGGACCAGTGCAGGCCGAGGCCAGAGAGACAGCTGGGCTTCCAGGTAGGGTCAAAGTG
AGGTGGGCAGCAGGTGTGGGGGCCAGGGGACTCGGGGACTTCCTCTCCGGCTGGGCCCG
CCTGACGTGGGAGGCAGCCAGGGTTAATCATTTCCACGAAGCCTTGACCCCACCTGCCT
TGGCGCTCTGCTCCCGCCTCCCACTGCCCCTCAGGCCAGCTCAGGAGCCATGGGGCGCT
GGGCCTGGGTCCCCAGCCCCT

ggttcactcacccctcctcccacctcggcagccctgggatgtcgctgctgactcaggag
gaacccgaggtgccgtagcggctgctccaatattgcagaagaggttcctcaggcagctc
tgcccacagccccaagtcacgaattccgtgactccagctccatcccaggccccagggta
cctggcccagggttgtgctgccgcagacttggcctgtaccatccaggcggcggtgggga
gctggggttggaaaggcttcttggagtggactcctgggtctgtctgggagacggggagg
aagggacactctgaacatcaccaggggctgctggggggccctggccacccccagagtca
gaacaggcaggtggggcaggatctcaggtcatcctatgctacactcagccattgcgtgg
cccctctcctccctgtgcctggccttttggccagccctggggccaccgagaggatgcag
caccgaaccctccaggagcccccagtgctgccgtctgtgggacagggacaatcccatcc
ccaCTGCTACTGTCTGTGCTGTGCTGGGC[ga]CAGAGCTGGACACCTCCAAGGCCCAG
cgcccgtagtggctctcatcatggacaattcacaggcagatggtggccagctctgtggc
ctgcagggactgggagcggcgccagaccatctaggccccaacctatctgcattatcctg
gaagacttcctggaggaggcttctaagctgaggcccaaggaccatgtcaggtctaggac
taggaccagtgcaggccgaggccagagagacagctgggcttccaggtagggtcaaagtg
aggtgggcagcaggtgtgggggccaggggactcggggacttcctctccggctgggcccg
cctgacgtgggaggcagccagggttaatcatttccacgaagccttgaccccacctgcct
tggcgctctgctcccgcctcccactgcccctcaggccagctcaggagccatggggcgct
gggcctgggtccccagcccct

FIGURE 23j

41365
TTTATGGCACAAATGGGGCCGGGGGCAGGCCCAGGGGCAATTCAACAGGAGGCAAGAGC
CCAGGGCTCCAGAGTGGAGAGACAGGAGGCAGCTCAGTCCCCAGACCCCAGCAGAGCAT
CTGGGGCCTCGGCCCCACTCCAGAGCTTCTTCCTGAGGGAGCCATGCACAGCAATGCTG
GGAGAGGGACTGATGGGGTGGGGTCAGGCCTCCTGCCACAGAGCTGGGCTGCAGAGCCC
AGATGGAAAGACACAGTGAAGAGCTCAACCTCCTTCCAAGCTCTCCTTCTCAGGGCTTC
AGGTTCCAGAGCCCCAGGGGAGCTCCCAGCCAGGGGCAGGGTCACCTTGATATTCACAA
CTGGGCTTGTGGGGGCCATCTTCAGTGCAACCGTTGTGACAAAGTCAAGAGGCTGCCTC
CCTGAAGCAGACCCACTGCCTACGCCACACTGACGGTCCAGAGGCCCCCTCCTGAGGGC
GGCCAGCAAGGGGCACTGTGGCAGCTCCCACTGTGCCTGTCCCAGACTGGGTCAGCAGG
TCTCTCTGGACAGCACACTGCACCAAGTA[tc]GCCCACCAAAAACGCATCAGGTGTGG
CCATGGCCCACAGTACCTTCTTCATTCCCTGCCTCTAACATGTGCGGTCTGAATGAATT
TTGTCACTCTTCTGCCATTTATAAAGGAGAAGACAGTGATCCAAAGCTATGCATGTTTC
TGAAGCCCTCAAGGAAGCTCGGTGCAGGCCATCACTTCTTTTGGCAGAAGGCGGGCTGT
GGTCTCTATGTACACACGCGAGCCCGCCAGTGACGTGCGGCAGTGCGTGGCGTCCAGGC
TGGGACAGGGGCCTTTCAAGTCTCCCCAGGGACCGGTGTTTTCTACAACAGACAGGTGC
TCCCAGACCGTTGGGGTACAGGCCAGGCCGTCTACACCACAGTATTGAGGGAGCTGCGG
CTGTGGCGGCCACCCCCTGGCAGTGCCTCTGCAGCTGGGGTGCTCCCGCTCTGGGCAGG
GTCAGGGGGCACGAGCAGGGC

tttatggcacaaatggggccgggggcaggcccaggggcaattcaacaggaggcaagagc
ccagggctccagagtggagagacaggaggcagctcagtccccagaccccagcagagcat
ctggggcctcggccccactccagagcttcttcctgagggagccatgcacagcaatgctg
ggagagggactgatggggtggggtcaggcctcctgccacagagctgggctgcagagccc
agatggaaagacacagtgaagagctcaacctccttccaagctctccttctcagggcttc
aggttccagagccccaggggagctcccagccaggggcagggtcaccttgatattcacaa
ctgggcttgtgggggccatcttcagtgcaaccgttgtgacaaagtcaagaggctgcctc
cctgaagcagacccactgcctacgccacactgacggtccagaggccccctcctgagggc
ggccagcaaggggcactgtggcagctcccactgtgcctgtcccagactgggtcagcagg
tcTCTCTGGACAGCACACTGCACCAAGTA[tc]GCCCACCAAAAACGCATCAGGTGTGG
Ccatggcccacagtaccttcttcattccctgcctctaacatgtgcggtctgaatgaatt
ttgtcactcttctgccatttataaaggagaagacagtgatccaaagctatgcatgtttc
tgaagccctcaaggaagctcggtgcaggccatcacttcttttggcagaaggcgggctgt
ggtctctatgtacacacgcgagcccgccagtgacgtgcggcagtgcgtggcgtccaggc
tgggacaggggcctttcaagtctccccagggaccggtgttttctacaacagacaggtgc
tcccagaccgttggggtacaggccaggccgtctacaccacagtattgagggagctgcgg
ctgtggcggccacccccctggcagtgcctctgcagctggggtgctcccgctctgggcagg
gtcagggggcacgagcagggc

FIGURE 23k

41408
TTAATATAAATAGGATATCATAATAAATAGAAATCATGCCAGGTCAGACGCACAGCACG
CTTGGAGCTCAGGGTTCCCTGAGACCCTGACCCTAAGTTCTGCTGTTCCCTTGCCCTGG
GGACCAGAGACGGCCTCCAGTCCCCCTCAAGTACCTCTGTGTGACCTCACAAGGCCTCC
CAGGGCCTCAGATGTGAGCTGCTACTCTGAGCTACCCCAGCCCCTTCTTACAGACCTTT
ACCCAGAGGAAGAGCCTGGGTCCCTCAGAACCTCTGCACCTGACTTAGCAACCTGCCCC
TGCCCTACCCACCTCCACAAACCCCTGCTGCAGGTCCAGCCATCAGACCCTGGCCATCC
CAGGCTGCAGGGAAGATCACGGGGAAGAGAACGAAGAACCTACCAAAGCTTTCCAGGCC
TCTCCTCCTCCCAGTGTCTTCCTTCCCAGGCCTGAAGGTGGCTTCTCTGCCTCCCCAAG
AGCCTGAATGCCAAGTGACCTCCTTCTGGAAACTTCTGCCAGATTGTTCCTATGCCCAA
GTTCTCTGATCATCCTCAAAAGAAGACAG[ac]CTTCCATCCCAGAGGCCCCTCTCTAT
CTTCCACTCATCAAACTTCTAGGGGACAAGGAGTCCTTTGGGATCCTAGCCCCTCTGGC
CCACCTAAGTCCCAACCTAAGGGGCAGCAAAGGCACAGATGGTGATAATTTGCTGGGGG
CTGGTCCACTCCCCTGGGCCCTGCTGTCTCACCCTGTGGTCAGGGCTCTTGTAGATGAC
TTGTGTAGTTTGTTCACTGCACAAAGTGAGCAAGGGGCCAAAGGGACAAGTAGAGGCAG
AAGTCCAGCCCACGCTCCCCAGTCCACAATCTCCCAGAGGAAGGGGCACCTTCTTCTAG
CTCCCTCCCTATGGAAGTTTCCACTCTGCTCAGCTTCATCACAGCCCAGCCCAGAGTGG
AGTGGACTGGCCAGGCACCCTCGGGGTCTGCCAGCAGCCCCCATTTGGGTTTAGCGATG
CCCTGGGCCCCAGCCACCCTT

ttaatataaataggatatcataataaatagaaatcatgccaggtcagacgcacagcacg
cttggagctcagggttccctgagaccctgaccctaagttctgctgttcccttgccctgg
ggaccagagacggcctccagtccccctcaagtacctctgtgtgacctcacaaggcctcc
cagggcctcagatgtgagctgctactctgagctaccccagccccttcttacagaccttt
acccagaggaagagcctgggtccctcagaacctctgcacctgacttagcaacctgcccc
tgccctacccacctccacaaacccctgctgcaggtccagccatcagaccctggccatcc
caggctgcagggaagatcacggggaagagaacgaagaacctaccaaagctttccaggcc
tctcctcctcccagtgtcttccttcccaggcctgaaggtggcttctctgcctccccaag
agcctgaatgccaagtgacctccttctggaaacttctgccagattgttcctatgcCCAA
GTTCTCTGATCATCCTCAAAAGAAGACAG[ac]CTTCCATCCCAGAGGCCCCTCTCTAT
CTTCCACtcatcaaacttctaggggacaaggagtcctttgggatcctagcccctctggc
ccacctaagtcccaacctaaggggcagcaaaggcacagatggtgataatttgctggggg
ctggtccactcccctgggccctgctgtctcaccctgtggtcagggctcttgtagatgac
ttgtgtagtttgttcactgcacaaagtgagcaaggggccaaagggacaagtagaggcag
aagtccagcccacgctccccagtccacaatctcccagaggaaggggcaccttcttctag
ctccctccctatggaagtttccactctgctcagcttcatcacagcccagcccagagtgg
agtggactggccaggcaccctcggggtctgccagcagcccccatttgggtttagcgatg
ccctggccccagccaccctt

FIGURE 231

41409
TGTGACAATCAGCAAAGCCCCACCCAGGCCCCCATCTGGGATGATGGGAGAGCTCTGGC
AGATGTCCCAATCCTGGAGGTCATCCATTAGGAATTAAATTCTCCAGCCTCACTCTCGG
CTCTTTCCTACTTGTTAGTAGTCTTGGGATGGTGGTAGTCAGAGGCAGGGACTGAAGAG
GTGAGGGAATGACAGAACCGACATTTACCAGGCACCAGCTGTATACATTACACATGCCA
TCTCCTTTAATCTGCATCACAACCCTGTGAGATCAGTGCTATTCTTAGACCCATTTCAC
AGGTGAGCGAACTGAGGCCTTTAAAAGGTTACATCAACCTCTCAAGATCAGACACCAAA
CCATAGTTCAGCTAGGTGTCGCAGGGGGGAATACTTATTAAGTGCTAAGCACTGTATAT
GTATTGGTTCACTTAATCCTCAACAACCCTATGAGGTAGCTCCTGTTTAGAGACCCCCT
TTTTTTAGAGGAGGAAACTAAGGCTTAGAGTGCAAGAGGGAGGTCCTTTGCGCAAAGGC
ATGGAGGAGATTTGAATTTAGGTTTAGGG[ac]TGGGCCAGGAAGGGCACGGCAGCCGT
TAAAAAAAGAGGCCCCCCTGGGAGGAGGGGAGCTGAAAGCCCTCTCCAACACCCACCCC
AATCCTGGATTCAGACACAGACATTTCTGTGACATCCCTAACTTCCCACCTGCTACCTC
AGGCCACAGCACCCAGGCACTAGGGCTCCCCTAGGCAGGTTTTTGAGGCATGTATTATT
TTTGCAACACGGACATACATGTACCTCCTCCTGGTACTGCCTGGGGCTGCTGCAATAAG
TTACCCTTTCCCCATTCTCATCTGTATGTGAAGTTCCCTGGCAAGGCCAAAGCCCAGGG
CATCAGAATGAGCTTCCTGAACACCACATCCAGGCATAGAAGAGTTGTGTCATACATAG
CTCAAGGTTACCCAGAACAGCAGGAGATGTGGTCCAGCATTTGGGCCTTGAGATCCCCC
CATTCATCCTCTTGATTGTCC

tgtgacaatcagcaaagccccacccaggcccccatctgggatgatgggagagctctggc
agatgtcccaatcctggaggtcatccattaggaattaaattctccagcctcactctcgg
ctctttcctacttgttagtagtcttgggatggtggtagtcagaggcagggactgaagag
gtgagggaatgacagaaccgacatttaccaggcaccagctgtatacattacacatgcca
tctcctttaatctgcatcacaaccctgtgagatcagtgctattcttagacccatttcac
aggtgagcgaactgaggcctttaaaaggttacatcaacctctcaagatcagacaccaaa
ccatagttcagctaggtgtcgcaggggggaatacttattaagtgctaagcactgtatat
gtattggttcacttaatcctcaacaaccctatgaggtagctcctgtttagagacccccct
tttttagaggaggaaactaaggcttagagtgcaagagggaggtcctttgcgcaaaggc
atggAGGAGATTTGAATTTAGGTTTAGGG[ac]TGGGCCAGGAAGGGCACGGCAGCCGt
taaaaaaagaggcccccctgggaggaggggagctgaaagccctctccaacacccacccc
aatcctggattcagacacagacatttctgtgacatccctaacttcccacctgctacctc
aggccacagcacccaggcactagggctcccctaggcaggtttttgaggcatgtattatt
tttgcaacacggacatacatgtacctcctcctggtactgcctggggctgctgcaataag
ttaccctttccccattctcatctgtatgtgaagttccctggcaaggccaaagcccaggg
catcagaatgagcttcctgaacaccacatccaggcatagaagagttgtgtcatacatag
ctcaaggttacccagaacagcaggagatgtggtccagcatttgggccttgagatccccc
cattcatcctcttgattgtcc

FIGURE 23m

41420
CCACCACCGAGGCCGAGCTGCTGGTGTCGGGCGACGAGAACTGCGCCTACTTCGAGGTG
TCGGCCAAGAAGAACACCAACGTGGACGAGATGTTCTACGTGCTCTTCAGCATGGCCAA
GCTGCCACACGAGATGAGCCCCGCCCTGCATCGCAAGATCTCCGTGCAGTACGGTGACG
CCTTCCACCCCAGGCCCTTCTGCATGCGCCGCGTCAAGGAGATGGACGCCTATGGCATG
GTCTCGCCCTTCGCCCGCCGCCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGCCAA
GGTCCTTCGGGAAGGCCAGGCCCGTGAGAGGGACAAGTGCACCATCCAGTGAGCGAGGG
ATGCTGGGGCGGGGCTTGGCCAGTGCCTTCAGGGAGGTGGCCCCAGATGCCCACTGTGC
GCATCTCCCCACCGAGGCCCCGGCAGCAGTCTTGTTCACAGACCTTAGGCACCAGACTG
GAGGCCCCCGGGCGCTGGCCTCCGCACATTCGTCTGCCTTCTCACAGCTTTCCTGAGTC
CGCTTGTCCACAGCTCCTTGGTGGTTTCA[gt]CTCCTCTGTGGGAGGACACATCTCTG
CAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCCTGGGGTTGGAAGAA
ATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCCTGGGACAAGCCT
CAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTCCTGGCATAACA
CCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTCTCCCAGCCCTGCAACTCTTACG
CTCTGGTTCAGCTGCCTCTGCACCCCCTCCCACCCCCAGCACACACACAAGTTGGCCCC
CAGCTGCGCCTGACATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCT
CCTAGACCACGCCCACCACTTAGACCACGCCCACCTCCTGACCGCGTTCCTCAGCCTCC
TCTCCTAGGTCCCTCCGCCCG

ccaccaccgaggccgagctgctggtgtcgggcgacgagaactgcgcctacttcgaggtg
tcggccaagaagaacaccaacgtggacgagatgttctacgtgctcttcagcatggccaa
gctgccacacgagatgagccccgccctgcatcgcaagatctccgtgcagtacggtgacg
ccttccaccccaggcccttctgcatgcgccgcgtcaaggagatggacgcctatggcatg
gtctcgcccttcgcccgccgccccagcgtcaacagtgacctcaagtacatcaaggccaa
ggtccttcgggaaggccaggcccgtgagagggacaagtgcaccatccagtgagcgaggg
atgctggggcggggcttggccagtgccttcagggaggtggccccagatgcccactgtgc
gcatctccccaccgaggccccggcagcagtcttgttcacagaccttaggcaccagactg
gaggcccccgggcgctggcctccgcacattcgtctgccttctcacagctttcctgagtC
CGCTTGTCCACAGCTCCTTGGTGGTTTCA[gt]CTCCTCTGTGGGAGGACACATCTCTG
CAGCctcaagagttaggcagagactcaagttacaccttcctctcctggggttggaagaa
atgttgatgccagaggggtgaggattgctgcgtcatatggagcctcctgggacaagcct
caggatgaaaaggacacagaaggccagatgagaaaggtctcctctctcctggcataaca
cccagcttggtttgggtggcagctgggagaacttctctcccagccctgcaactcttacg
ctctggttcagctgcctctgcaccccctcccacccccagcacacacacaagttggcccc
cagctgcgcctgacattgagccagtggactctgtgtctgaagggggcgtggccacacct
cctagaccacgcccaccacttagaccacgcccacctcctgaccgcgttcctcagcctcc
tctcctaggtccctccgcccg

FIGURE 23n

41452
GCCTATGGTGCAGGGCTGGCAGAGGCGGGGCCAGGATTCTAGCTTCCCCACACACCAGC
CCTGTGGCATCATTCTTCCCAACGTCCAAACGTTTTTCCAAGGGGGAGAAATGGACTGG
GTCATGTAAAGAAATACTCATTTTTAGGGCTTTTTATGTGGCCTTCAAAGCACGTTGCA
AACAAATCCCTTTCACTCCTCAGAGGAGGAGCCATTAGGAAGGTAGGGGGCGACAGGCA
CAGCCTACAGCCTCTCCTCAGGAGGACAGAGGGGGTCATCGCATTTGAGCCCCCTGCAG
TCATCTCGGGGGCTCCTGAGGGTCCAGGTCCACATGTTCGAGGGTCTGCAGCACATCCA
CGGCGCTGTAGGACTTCCAGGCCTGCATGTTACAGCTCTTCAGGATGGCTCCCAGCTGC
CTGCCAGGGCCTACTTCGAAAGTTTGGGGGAACCCCCTGCCCTTTTTCCTTTCGTATAT
GGCATGCATCGTCTGCTCCCACTTCACTGGGGAGACCAGCTGCTGGGCCAGCAGCTTGT
GGATGTGCCCGGGATGCCTGTATCTATGC[cg]CGTGGACGTTGGAGTAGACAGAAACC
AGAGGCTTCTTAATGTCGACTGCCTTTAAAGCTTGCGTCAGGGGCTCCACGGCTGGCTC
CATGAGGCGGGTGTGGAATGCGCCACTAACCGGCAACATCCTGGTGCGTCTGAAATGAA
ACTTAGAGGAATTCTTCTGGAGAAACCGTAGAGCCTGGGGAAGGAAGGAGGTTTCAGCC
GAGCAATGTCCCAGAAATCCGCCTTTACAGATCTGACCATTCACAGGGCCAAACTGGGA
GGGTGACCACAAAGAGACCCACAGCTGCTAGATGTGGACATGTGACCTGTCTGTCCCAG
CACCATCCCCAGGCAATTCACTTAACATCCTGGAATCTCTTCTGTCCCAGCCTTCAAAT
AAGCACAGTTCCATCTACTTCACAACGCTGCCAGGAAGAGCAAACCCTACAAGGCATGC
AACAGTGTCTGGTAGAGGAAA

gcctatggtgcagggctggcagaggcggggccaggattctagcttccccacacaccagc
cctgtggcatcattcttcccaacgtccaaacgtttttccaagggggagaaatggactgg
gtcatgtaaagaaatactcatttttagggctttttatgtggccttcaaagcacgttgca
aacaaatccctttcactcctcagaggaggagccattaggaaggtagggggcgacaggca
cagcctacagcctctcctcaggaggacagagggggtcatcgcatttgagccccctgcag
tcatctcgggggctcctgagggtccaggtccacatgttcgagggtctgcagcacatcca
cggcgctgtaggacttccaggcctgcatgttacagctcttcaggatggctcccagctgc
ctgccagggcctacttcgaaagtttgggggaaccccctgccctttttcctttcgtatat
ggcatgcatcgtctgctcccacttcactggggagaccagctgctgggccagcagcTTGT
GGATGTGCCCGGGATGCCTGTATCTATGC[cg]CGTGGACGTTGGAGTAGACAGAAACC
AGAGGCTtcttaatgtcgactgcctttaaagcttgcgtcaggggctccacggctggctc
catgaggcgggtgtggaatgcgccactaaccggcaacatcctggtgcgtctgaaatgaa
acttagaggaattcttctggagaaaccgtagagcctggggaaggaaggaggtttcagcc
gagcaatgtcccagaaatccgcctttacagatctgaccattcacagggccaaactggga
gggtgaccacaaagagacccacagctgctagatgtggacatgtgacctgtctgtcccag
caccatccccaggcaattcacttaacatcctggaatctcttctgtcccagccttcaaat
aagcacagttccatctacttcacaacgctgccaggaagagcaaaccctacaaggcatgc
aacagtgtctggtagaggaaa

FIGURE 23o

41475
ACGTTATCAGGCACAAACCCCCTCCAGACACCTGAGCCTCCCCCACAGGCTCCCAGTGA
GGAGCCATCACATGCCCAGGCCAGCCGAGGGGCCCTCAGGCATGGGGATCTGGGCAATG
GCAGCAAGCTGGGCGGGGGGTGCAGCCAGGATGACAGCAGATCTGCAGGGCGGGGTCCT
CGCCCCGGGCCACCTGGCTGGGGCCGAAGGTCACAGCTGCGTCTAACTGGGCCTTGAGC
AGCTGAAGCTGTTTCAGGGCTTGCAGCACCTCTGGGGTGGCCCCGGCCACACCCCCCAG
CAGGTTGTAGTTCTCACCAGGGTCCTTGGACAGGTCATAGAGCAGCGGGGGCTCATGAG
CAGTCAGAGAGCTGGAGGCGTGGCAGGCAGGGTCTGCAGTGGTATCACTGTGGGCAGAG
CCTGGGGAGGGGGCCAATTCTGTGCACAGGGCAAGGGCGAGAGGAGGGGCCAGGGATCT
AGGGCTCCGGGGAGGGGTCAGCAGGTCGGGGGGAGGGATCCACGGGGAGGGGTTACCCT
GGGTGAAGAAGTGAGCCTTGTACTTTCCA[cg]TCCGCACAGCAAAAACCCCACGGACC
TCGTCTGGGTAGGACGGGTAGAAGAAGAGAGACTGCCGAGGGCTCTGGGGGCAGAGTCA
GGGGTCACGGGGCGGGGCAGGCCCCAAGCACTGCACATACCTGGGGCTGCCAGCCCTGG
TGGGAGGCCCTGGACGTGCACCGCTTCTTGCCCACCCAGGAACCTGAGAGGTGGCGCCA
CTTGGATGCCACTCAGTGCAGGAGGCACTGAGGCACAGACTCTCAGGCACTGCCCACAC
TCACCCCAGGGGAAGGCCAGGACAGGGGCCAAGGATCTGGGATCAGGGGTCACCGGCCC
TACCTTGCCTGTGCCCAGCAGCAGGGGGCTGAGGTCAAAGCCATCCAAGGTGACATTGG
GCAGTGGGGCCCCAGCCAGGGCTGCCAGGGTAGGCAGCAGGTCCAGGGAGCTGGCCAGC
TCGTGGGTCACGCCTGGGGGC

acgttatcaggcacaaaccccctccagacacctgagcctcccccacaggctcccagtga
ggagccatcacatgcccaggccagccgaggggccctcaggcatggggatctgggcaatg
gcagcaagctgggcggggggtgcagccaggatgacagcagatctgcagggcggggtcct
cgccccgggccacctggctggggccgaaggtcacagctgcgtctaactgggccttgagc
agctgaagctgtttcagggcttgcagcacctctggggtggccccggccacaccccccag
caggttgtagttctcaccagggtccttggacaggtcatagagcagcgggggctcatgag
cagtcagagagctggaggcgtggcaggcagggtctgcagtggtatcactgtgggcagag
cctggggagggggccaattctgtgcacagggcaagggcgagaggaggggccagggatct
agggctccggggaggggtcagcaggtcggggggagggatccacggggaggggttaccct
gggTGAAGAAGTGAGCCTTGTACTTTCCA[cg]TCCGCACAGCAAAAACCCCACGGACC
tcgtctgggtaggacgggtagaagaagagagactgccgagggctctgggggcagagtca
ggggtcacggggcggggcaggccccaagcactgcacatacctggggctgccagccctgg
tgggaggccctggacgtgcaccgcttcttgcccacccaggaacctgagaggtggcgcca
cttggatgccactcagtgcaggaggcactgaggcacagactctcaggcactgcccacac
tcaccccaggggaaggccaggacaggggccaaggatctgggatcaggggtcaccggccc
taccttgcctgtgcccagcagcagggggctgaggtcaaagccatccaaggtgacattgg
gcagtggggccccagccagggctgccagggtaggcagcaggtccagggagctggccagc
tcgtgggtcacgcctgggggc

FIGURE 23p

41541
GAATGGTAAGAAACATTCTTCAGCTCAAGATGGTGACCAGAGGCATCCAGCACTCACTT
CCTTCACAAAGGACTCAAACAGCAAATGAATAATCACATGTCAAGTAGAGCAGCTTAGA
AAGAACACTGGAATTCAGAGGGAAAGGACAAGGAACTTCGGAAACATGCAAAGAGAATG
ATGTGAAGCAGCCGGCCCAGCCAGGATCAGCTCAGATCCAAGAGAAACTGCCCAACGTA
GGGAAAAGGTAAATGAGAGATCCCCGCAAGGCTGCATTCCCACCACAGACTCCTGTGGC
CCTAGCCACAGAGAGCCCCTTGGCCCTCATGGGCTTTGAGACTAGTATAGAGAGCCGCC
TGCATTGTTCCAAAGAGGGATTTTATGATGGGTCCTACACATCCTCTGAGACCTGAGCA
GCTGCAGCACAGCACCATTTTGAGAGCCCACCCCTGACCAGACCCCATCCCGCCCTGGG
GCTCAACAGCCCCTGCATCTCCACATCCATGGAGTCCTGCTGACATTCCGCCATGTCCA
CCCAGAAGGCTGCAGCCTCACAATGCAGG[ag]TGACTGGGTCCCCAGCAATCTAGTCT
ACACATGTCCTATAACCTGGGAATGGGTGGTGCACCACACCAGGGAGGCTGCCCCTGGG
ACAAAGGGAGCCAAAGCCCATGTTTCCCAGAGCCGCAGAGCTGCCCGCCTGGGACCACT
GCCACTGACAGCACCCCCACCATCCCCCCAGCAGCGGGGTCACTGTGCACTTGTGATAT
GGTTTGGCTGTGTCCCCACCCAAATCTCATCTCCAGTTGTAATCCAAATTGTAATCCCC
ACGTGTCAGGGGAGGGACCTGGTGGGAGGTCATTGGATTACAGGGGCGGTTTCCTCCAT
GTTGTTCTCATGATAGTGAGTAAATTCTCATGAGATCTGATGGTTTTATAAGTGTTTGA
TAGTTCCTTCTTCACACACACTCTCTCCTGTCGCCATGTGAAAATGTCCTTGCTTCCCC
TTTGCCTTCCGCCATGACTGT

gaatggtaagaaacattcttcagctcaagatggtgaccagaggcatccagcactcactt
ccttcacaaaggactcaaacagcaaatgaataatcacatgtcaagtagagcagcttaga
aagaacactggaattcagagggaaaggacaaggaacttcggaaacatgcaaagagaatg
atgtgaagcagccggcccagccaggatcagctcagatccaagagaaactgcccaacgta
gggaaaaggtaaatgagagatccccgcaaggctgcattcccaccacagactcctgtggc
cctagccacagagagccccttggccctcatgggctttgagactagtatagagagccgcc
tgcattgttccaaagagggattttatgatgggtcctacacatcctctgagacctgagca
gctgcagcacagcaccattttgagagcccacccctgaccagaccccatcccgccctggg
gctcaacagcccctgcatctccacatccatggagtcctgctgacattccgccaTGTCCA
CCCAGAAGGCTGCAGCCTCACAATGCAGG[ag]TGACTGGGTCCCCAGCAATCTAGTCT
ACACATGTCctataacctgggaatgggtggtgcaccacaccagggaggctgcccctggg
acaaagggagccaaagcccatgtttcccagagccgcagagctgcccgcctgggaccact
gccactgacagcacccccaccatccccccagcagcggggtcactgtgcacttgtgatat
ggtttggctgtgtccccacccaaatctcatctccagttgtaatccaaattgtaatcccc
acgtgtcaggggagggacctggtgggaggtcattggattacaggggcggtttcctccat
gttgttctcatgatagtgagtaaattctcatgagatctgatggttttataagtgtttga
tagttccttcttcacacacactctctcctgtcgccatgtgaaaatgtccttgcttcccc
tttgccttccgccatgactgt

FIGURE 23q

41646

CACCTCCCTTAACTCCCCAGCCATGCCCCGTGGGTATCTGTTTTCCCAGTTTTGTAGAT
GAAAGCACAGCTCAGAGAGGTTTACTCAGTTGCCTGGAGTCACACAGTCAACAAGTGGA
GAGCCAGTCATTGAATCTGGTACCACAAACTCTTCCTGCTGCAACAGCTGTGCTTTTGC
AGGCACTGACTTTGGAATACCCTCAGCTGATTCACAGGGTCCTTTGTCCTGGGGAATGG
CCTTCCCTGTCTCCTTCAGGGAAAGGGTTTCATCCTTCAGGGAAGATTCATTGAATCAG
GATTTGCTGGGTTTTTTCATTTTTTTTTTTCATTTCTTTTTTTTTTACACGAATGGGC
TTCCTGGCCCGCATTTTGATTTGCGCTTGGGTTTATGAATTGAGGAATCACAGTCAGCC
TTGGGAATTAGTTGCAAGATAAATATTGCAATCCTGGTTAAGGACTTAAGAATTGTCAC
TTGTGTGTGTATATTGTTGTTGTTGTTGCAACGGT [gt] CTGTGTACGCACGGTTACAG
TGGATCAAATTTGGGGAGTTAGGAAGTGGCGTTGGTTTGTGGTTAGACTTGGGGGAGGT
GTCGCTTTCGGTTGTTGGTGTGCTGGTGGCTGTGTTCCTGTGATATGGAATGTACTGTC
TGAGAATGTGTTCAGGGGTCTGTGGTTATGTGGATATGGGTGTGTAGCTGCTGATGACA
TGGATGGAGGGATGTATCTGGGTGTGTTTCTGCAGAACAAGTGATACCTGTACCATGTG
ACTTTGTCAGTTCCACCATGTCCAGGCACAGGTCGGGGGGGTTGTCCATGGTTCTGAAC
GTATCTGCCCCCATTTTACAGATAGGAAACCAAGACTTAGAGAGGCCAAGTCATCTGCT
TGAAGTCATCTAGCTGAGAAGCGGCTGAGCCTGAAGGGAAACCAGGGCTGCCTTCAGAG
TCCAGCCTCTTTTCCCTGCTCCCCAGGAAAGGTTTTAGTAACAATAAAAGGTTTAAATG
CCAGCAAAAGGTCTAAACGCC

cacctcccttaactccccagccatgccccgtgggtatctgttttcccagttttgtagat
gaaagcacagctcagagaggtttactcagttgcctggagtcacacagtcaacaagtgga
gagccagtcattgaatctggtaccacaaactcttcctgctgcaacagctgtgcttttgc
aggcactgactttggaataccctcagctgattcacagggtcctttgtcctggggaatgg
ccttccctgtctccttcagggaaagggtttcatccttcagggaagattcattgaatcag
gatttgctgggttttttcatttttttttttcatttcttttttttttacacgaatgggc
ttcctggcccgcattttgatttgcgcttgggtttatgaattgaggaatcacagtcagcc
ttgggaattagttgcaagataaatattgcaatcctggttaaggacttaagaattgtcaC
TTGTGTGTGTATATTGTTGTTGTTGTTGCAACGGT [gt] CTGTGTACGCACGGTTACAG
TGGATCAAATTTGGGGagttaggaagtggcgttggtttgtggttagacttgggggaggt
gtcgctttcggttgttggtgtgctggtggctgtgttcctgtgatatggaatgtactgtc
tgagaatgtgttcaggggtctgtggttatgtggatatgggtgtgtagctgctgatgaca
tggatggagggatgtatctgggtgtgtttctgcagaacaagtgatacctgtaccatgtg
actttgtcagttccaccatgtccaggcacaggtcgggggggttgtccatggttctgaac
gtatctgcccccattttacagataggaaaccaagacttagagaggccaagtcatctgct
tgaagtcatctagctgagaagcggctgagcctgaagggaaaccagggctgccttcagag
tccagcctcttttccctgctccccaggaaaggttttagtaacaataaaaggtttaaatg
ccagcaaaaggtctaaacgcc

FIGURE 23r

44901
TTGCCCCACAGACAAGATGATCCCCCCTGGCATGTTGTTAGGGGCAAATTGCTGTCCTG
CTCAGAGTGGCATCTTTCAATGTTGCCTCCATCTTGGCCAAGAGGTCCCTGCCTCCTGA
TCCGGCACAGCTGAGCTGAGGCAGATGTGACCAGTTTTCAAGCTACCAGCCCTGGGCAG
AGGAAGATGTCAACAATTCCAGAGCAGAGGGAAGAGGCACCTTCCTTGACCACACCAGT
GGCCTCCTGAAGTTCCATGCTTTTAAGAGCTGGGACCTTGGGAGGATGATTCAAACCCT
CAATTCCTCCTCCCTGGGAACTTTTTACCACCTTTACCTATTTATCAAAATCATATTCA
TCTTTACCATCACTGTCACTGTAATCTACATTCCATCACCTTTATCAGGTGCTGCTGAG
TACAAAGCACTTGGGATGGGAGACACAGCACTGAATTCACAAACATTGGACCAAACTGT
TTGTCCCCATCTGGGTTCATGAGGCCACCTCTTTGCTCAATCCATGCCTCTTGCCCTCA
GTCAACAAGACATTCCTAGAGGGAAAGGG[ct]TGCTGCTCTGGGAGTCAACCTGAGTT
CCTCCCTCCTGGGAAGCTGGGTTGGCAAGATTCTAGGACACTCACCTGCATGGACATCA
CCTCTGTGACAAATGCTTACCTGTTTCTCATCTTCAGACTTGGCGATATCAAGCCTGTT
CTGGACCATGACCAGGCTGGCTCATATCTCTGGTTTAGAGAAACCTATGAATAACTGGG
GACAAACAGACTCTTTGGTAGCAGCAGACACATGTGATCCATCAAGATCAACCAAGGTT
GCAACTGGAGCGTCCACTGCCAGAGACCTTTGGCTCTTCAAGCTCGGGACAAAAAAGAA
GACTCTGTTGTCCCTTGGTAACCCAGTCCCTGCTTTTGTAGCTATCACAGCAGAAAGCA
ACTCTTCCTGAAGACCAAACACTCGTCATCCACATTCCTTGAATGGCCAATCCTTCCAT
CTGGAGGCCTGGCTCAGAAAG

ttgccccacagacaagatgatcccccctggcatgttgttaggggcaaattgctgtcctg
ctcagagtggcatctttcaatgttgcctccatcttggccaagaggtccctgcctcctga
tccggcacagctgagctgaggcagatgtgaccagttttcaagctaccagccctgggcag
aggaagatgtcaacaattccagagcagagggaagaggcaccttccttgaccacaccagt
ggcctcctgaagttccatgcttttaagagctgggaccttgggaggatgattcaaaccct
caattcctcctccctgggaactttttaccacctttacctatttatcaaaatcatattca
tctttaccatcactgtcactgtaatctacattccatcacctttatcaggtgctgctgag
tacaaagcacttgggatgggagacacagcactgaattcacaaacattggaccaaactgt
ttgtccccatctgggttcatgaggccacctctttgctcaatccatgcctcttgccctcA
GTCAACAAGACATTCCTAGAGGGAAAGGG[ct]TGCTGCTCTGGGAGTCAACCTGAGTT
CCTCcctcctgggaagctgggttggcaagattctaggacactcacctgcatggacatca
cctctgtgacaaatgcttacctgtttctcatcttcagacttggcgatatcaagcctgtt
ctggaccatgaccaggctggctcatatctctggtttagagaaacctatgaataactggg
gacaaacagactctttggtagcagcagacacatgtgatccatcaagatcaaccaaggtt
gcaactggagcgtccactgccagagacctttggctcttcaagctcgggacaaaaaagaa
gactctgttgtcccttggtaacccagtccctgcttttgtagctatcacagcagaaagca
actcttcctgaagaccaaacactcgtcatccacattccttgaatggccaatccttccat
ctggaggcctggctcagaaag

FIGURE 23s

44907
GTTTGATGGGACAAGATAGGACAGTGGTTAAGAGTGTGACCTCAGCAGCTGACTGCCTG
GGTGTAAAGCCTACCATGTGGTCAAGCACACGGGTGGCTCTACCACTTACCAACCATGT
GACCTTGGGCGGTTAACAGCCCTGTGACTCGGTTTCCCCATCTGAAAAGTGAGGATCAT
AGCAGTATCTACCTCCTGCGGTGGTCGGAAGGCAGAAAAGAATTGGCACATGTGAAAGT
ACTTAGCACAGGCTTGGTGCATAGCAAGTCCTGAGGAAATGTATTCACTGTCATCAGTT
TCACCCGCTTTGAAAGGCAGGCAAAGAAAGCACCTGACAAAACCTTTTGATCCCCCACG
CCTTGTCTCCCACACCCAGGACATTCCCCTGACTCCCATCTTCACGGACACCGTGATAT
TCCGGATTGCTCCGTGGATCATGACCCCCAACATCCTGCCTCCCGTGTCGGTGTTTGTG
TGCTGGTAAGGGGTGACCCCAGCCTGGAGAGGCAGCGTGGCAGAGTGGCCAAGGGCCGA
GTCAGATGGACATGAGTCTAGTTCCTGGC[ct]CCGTCACTTACCACTGTGTTACCTTG
AGCAACTCTCTTGGCCTCTCTGAAATGCCCACATCGTAGAGTCACTGTGAGAATTAAAT
GAGATGAAGCAGGCAAAGCATTTATCCAAGGCCCAGCACACAGGGTATGCTCTAAAAAT
AATAGCTGCCATTCTGTTCTCTTGCTTAACCCTCTACCAGGCAGTTAGCAACCTCCTAT
GCAGTGGAAATGCAGCTCATCTGACTCATTCATTAAACAGACTTTTATTGACCACCTAT
TATGAGCTAGGTCCACAACAGCAAGATGAGAACCAAGGGAAAAAGTGCCTGTGATTAGA
TGGCTAGCAACCCAAAAGGGACCCTTGGGGTCCTCACGTCCATCCCATCTTCATGCCAG
GCAGAGCTCTTCTTTGAAAATCTGTGGAGTCAGAGGTGTAAGGCATTGGGACAGGTGGG
GGTGAGAGTTCCCCCCCTCAT

gtttgatgggacaagataggacagtggttaagagtgtgacctcagcagctgactgcctg
ggtgtaaagcctaccatgtggtcaagcacacgggtggctctaccacttaccaaccatgt
gaccttgggcggttaacagccctgtgactcggtttccccatctgaaaagtgaggatcat
agcagtatctacctcctgcggtggtcggaaggcagaaaagaattggcacatgtgaaagt
acttagcacaggcttggtgcatagcaagtcctgaggaaatgtattcactgtcatcagtt
tcacccgctttgaaaggcaggcaaagaaagcacctgacaaaaccttttgatcccccacg
ccttgtctcccacacccaggacattcccctgactcccatcttcacggacaccgtgatat
tccggattgctccgtggatcatgacccccaacatcctgcctcccgtgtcggtgtttgtg
tgctggtaaggggtgaccccagcctggagaggcagcgtggcagagtggcCAAGGGCCGA
GTCAGATGGACATGAGTCTAGTTCCTGGC[ct]CCGTCACTTACCACTGTGTTACCTTG
AGCAACTCTCTTGgcctctctgaaatgcccacatcgtagagtcactgtgagaattaaat
gagatgaagcaggcaaagcatttatccaaggcccagcacacagggtatgctctaaaaat
aatagctgccattctgttctcttgcttaaccctctaccaggcagttagcaacctcctat
gcagtggaaatgcagctcatctgactcattcattaaacagacttttattgaccacctat
tatgagctaggtccacaacagcaagatgagaaccaagggaaaaagtgcctgtgattaga
tggctagcaacccaaaagggacccttggggtcctcacgtccatcccatcttcatgccag
gcagagctcttctttgaaaatctgtggagtcagaggtgtaaggcattgggacaggtggg
ggtgagagttccccccctcat

FIGURE 23t

44927
GCCCTGCCCTGTAGTGGCTTCTCAATGAATATGTAGTTGCCTTATTCTCACAAACACCA
GGCTTTCCTCACATCAGCACCCGGTGTGATAGGTAAGAGTGTGTGATACTAGAAACGTC
AGCTTATCCAAAAATGTATTTCTTTCTCTCATGAGAGCCTCGTGAGCTCTCCAGCTTGC
TGGAACTTTCTAAGACCTAACACTTGCCAAATTCCTTGCAGCAATTGTCTGGTTTGTGG
TACCACAATCGAACCCACCACCCTGACGTATTTGCTGCTCAGAACCACCGATCTTCCAA
GTTCTCATCACTCCAGTGCAGCTCCTGTGACAAAACCTTCCCCAACACCATTGAGCACA
AGAAGCACATCAAAGCAGAACATGCAGGTGGAGTTTGGGTACCGCCGGCAGAGAGCGGG
AGGGGCTTGATGGTGTAGCCTCCTGGGCCCCACCAGAAATCCCCACTTCTAATAGTCTA
GTGTGATGTGCAGTGGTCATTGCCTTTGTTCTGCCCCAGCGCACCTGTCCGTAGCAGCA
GCAGTCAGTAGCAGCAGCTTGAGTGGCAG[ct]GGTTCTCAAACCTGGAAGCGTAGCGC
AGTGTAAGCTCCCACCAGCCCTGAGTGAGAGCTTGTTGGGGCACCTGGAAGGGTGTCA
GCCTCAGTGGTAGGCAGGCCTGAGTGGAAATCCTGATTCCAGCACTTATCAGCTACATG
ACCTTGGCAAGTGACTTCCCTTTTCTGAGCCTGTTTCCTTCTCTCCAGGATGGCAGTTA
TTAAAACCTACTTTGCAGGTAAATTTGGTGATAATCACAACAGCTGTCAGTTACAGAGT
GTTTCCTATGTGCAAGACACCATGCTAAGCACCTCGTGTATATTTTCTCATTTCATTCT
CACAACATCCCTCTGAGCATCCAGGCAGTCTGGATCCAGATCTCATGCTCTTTACCACT
AGATTGTACAAATATACCATAGGTTATAAGATTCCTGGCACTTGGTAGATGCTTGCTAA
GTATTGGCCATCGCCCCAACC

gccctgccctgtagtggcttctcaatgaatatgtagttgccttattctcacaaacacca
ggctttcctcacatcagcacccggtgtgataggtaagagtgtgtgatactagaaacgtc
agcttatccaaaaatgtatttctttctctcatgagagcctcgtgagctctccagcttgc
tggaactttctaagacctaacacttgccaaattccttgcagcaattgtctggtttgtgg
taccacaatcgaacccaccaccctgacgtatttgctgctcagaaccaccgatcttccaa
gttctcatcactccagtgcagctcctgtgacaaaaccttccccaacaccattgagcaca
agaagcacatcaaagcagaacatgcaggtggagtttgggtaccgccggcagagagcggg
aggggcttgatggtgtagcctcctgggccccaccagaaatccccacttctaatagtcta
gtgtgatgtgcagtggtcattgcctttgttctgccccagcgcacctgtccgtagcagCA
GCAGTCAGTAGCAGCAGCTTGAGTGGCAG[ct]GGTTCTCAAACCTGGAAGCGTAGCGC
AGTGTaagctcccaccagccctgagtgagagcttgttggggcacctgggaagggtgtca
gcctcagtggtaggcaggcctgagtggaaatcctgattccagcacttatcagctacatg
accttggcaagtgacttcccttttctgagcctgtttccttctctccaggatggcagtta
ttaaaacctactttgcaggtaaatttggtgataatcacaacagctgtcagttacagagt
gtttcctatgtgcaagacaccatgctaagcacctcgtgtatattttctcatttcattct
cacaacatccctctgagcatccaggcagtctggatccagatctcatgctctttaccact
agattgtacaaatataccataggttataagattcctggcacttggtagatgcttgctaa
gtattggccatcgccccaacc

FIGURE 23u

45046
TAAATTTCTTACAAGGTCTCTTCTAGTTCTAATTTTTTAAAAAATGTTATGACCTCTGC
CCAGATTTTTTGTCTCACTGGAATTTTATGAAATCAAATAGTTTGTAAGTGGACCATTA
TAGGACTGTTTTGCCCAGTTCTTTGTTGTAAGGGTGTTTGACCGGTTGAATCATGGTAT
TTAAAAAATTCTTATACAACTCCAGATCTAATGGTAGGCTAAGTTGTGGTGATGCTTAT
ACTCAGTGATATTGGGTGTGTATTATAAGAATGAAGAGAGCGGAGAACAAACATAAACA
TTAATGTTAATGACAAACATTAACCCAAGTACAAGGTTAATGTTTAGTCAATATAGCAA
ACATGTAATTTACAAGATTAAAAATAATTAGGCTTGTGATAAAGTCAATGAATTTCCTA
CGTAATTGTAACATTAGACTGTTTTATTATTTGTCCTGACATTTTGCAGAATCCAAGAT
TAATTAAAGAAATGGTTTCAAGAAGAGGGTGAATACTATAAAAATAGACTTACCTTCCT
GAATTGAGGAATTCATCAGGAAAGCCTCA[ag]GTGTGCAAATGAGCCATCCTTCCAGA
GGGAAATTTCTTAGAATTATCCCACGATTTGAGCCAAAGCACTTCCGATAGAATTTTTA
ACCTCTAGTTGGTTCTGCTCCTTCCATTTTTACTAATTTTTAAGAAAATACTATGACTT
ATAATTGTATCTGGAATGATTATCAACTCCTTTTCATCCACTGACTTAAATTTGATTAT
AAATATGCTTTACATAAAGATCTAGACCTTATAATTTGAATTCAAGTGAATTGTTGTGA
CTAGCATGTAAATTATTATTATGGATTGTAAATCTTAACATAGGTAGTTCTGTGCCCTT
AAATTGATAAACCAGTTATCTCTTGTAATCATGTGTACTAAGATATACGTAGTAAAGTG
ATTGTATCAGTTTTTATCATAAGCAGTCATAGTTCAGATAGTTCAGAAGTTTAGTGTCT
GCTGTTTCTATTAGGAAAGTG

taaatttcttacaaggtctcttctagttctaattttttaaaaaatgttatgacctctgc
ccagattttttgtctcactggaattttatgaaatcaaatagtttgtaagtggaccatta
taggactgttttgcccagttctttgttgtaagggtgtttgaccggttgaatcatggtat
ttaaaaaattcttatacaactccagatctaatggtaggctaagttgtggtgatgcttat
actcagtgatattgggtgtgtattataagaatgaagagagcggagaacaaacataaaca
ttaatgttaatgacaaacattaacccaagtacaaggttaatgtttagtcaatatagcaa
acatgtaatttacaagattaaaaataattaggcttgtgataaagtcaatgaatttccta
cgtaattgtaacattagactgttttattatttgtcctgacattttgcagaatccaagat
taattaaagaaatggtttcaagaagagggtgaatactataaaaatagacttaccttcct
gAATTGAGGAATTCATCAGGAAAGCCTCA[ag]GTGTGCAAATGAGCCATCCTTCCAGA
GGgaaatttcttagaattatcccacgatttgagccaaagcacttccgatagaattttta
acctctagttggttctgctccttccatttttactaattttaagaaaatactatgactt
ataattgtatctggaatgattatcaactccttttcatccactgacttaaatttgattat
aaatatgctttacataaagatctagaccttataatttgaattcaagtgaattgttgtga
ctagcatgtaaattattattatggattgtaaatcttaacataggtagttctgtgccctt
aaattgataaaccagttatctcttgtaatcatgtgtactaagatatacgtagtaaagtg
attgtatcagtttttatcataagcagtcatagttcagatagttcagaagtttagtgtct
gctgtttctattaggaaagtg

FIGURE 23v

45076

ACTTGGTGACTTTCTCTGCACCAGGTGAGCCCCTAGTCTACACTGCACTGCACCCCCCC
CCCACCCCGGCGCACGCACACACACACACACACACACACACACACACACACACACAGGC
ATGCACAGGCCCTCCTGTGAGAGATAGCCCTAAGGAGGGAACCGTCCCTAGAGCCGGTC
CCCAGCCGCTCGGCACTTCCCGCCCACGCGCCCGGTCCCACAGTGCAGCGGACCCTCAC
TCACCCCGCGGATGTCCCAGTACCCCAGTGTCATGGACATGATGCTGGTTGGTGTCGAT
TCTGCAGACAGGCCTCAGCTGGGCTGAACTGCGACCTCCTCTGGGGTTCCCGGCACGCA
GGGGCTGGACCTAGCGCCAGACCCGCCCCCTCGGCCCCGCTGCGCCCGCCGATCTTCAA
GGTCGTCACTTCCAACCGGCCGATCTTCAAGGTCGTCACTTCCAACCAACAGGCGCGGG
AGGCACGGAGCAGGTTGCTGGATCCTCACTGGCTGGAAGGAGTAAGATCCACCGCCACC
TCCGAGTGTTCAGGGAGCAAGGTCCGGAA[cg]CACTAGGAGGGGCTCGGCCTCGCCAG
CTTCCGTAGCCCCGCCCCGCCCCGCTCCGCTTCGGACCTCTGCTGGGTCCCCAGGGACT
CGGCTGTGCGCGTGAGAGTAAAGCCAGATCGTAAGAGAAAAGTTCTTCCCCCGTTTCTT
CTTCTCCGGACGTCGCCCAGCCTTCTGCCTCTCGGCTGCCGAGTTCCCACAGGCTCTGG
GAGACTGAGGCTGCCAGGGTCAGACTAAAGAGAGGTCTCAGAGAGTTTAATTCAACACT
TCTTGGCTACTAAGTCTTAGAAGTCTGATGGTGTGCTCTCTCTGCTGAGTTGGGGAGCG
TGAATGGAGGCTATGTCACCGAAGCTGATAGAGCTCAGTCTCTGTTGCAGATGCTCCCG
ACCCTTTTGCATTGGGCCAGTTCCCCAGCTCTGAGACTGGGTCCAGGCTCAGGAAGTGG
CCTATGTGTCAAGGTGGATTC

acttggtgactttctctgcaccaggtgagcccctagtctacactgcactgcaccccccc
cccaccccggcgcacgcacacacacacacacacacacacacacacacacacacacaggc
atgcacaggccctcctgtgagagatagccctaaggagggaaccgtccctagagccggtc
cccagccgctcggcacttcccgcccacgcgcccggtcccacagtgcagcggaccctcac
tcaccccgcggatgtcccagtaccccagtgtcatggacatgatgctggttggtgtcgat
tctgcagacaggcctcagctgggctgaactgcgacctcctctggggttcccggcacgca
ggggctggacctagcgccagacccgccccctcggccccgctgcgcccgccgatcttcaa
ggtcgtcacttccaaccggccgatcttcaaggtcgtcacttccaaccaacaggcgcggg
aggcacggagcaggttgctggatcctcactggctggaaggagtaagatccaccgccacc
tccgagTGTTCAGGGAGCAAGGTCCGGAA[cg]CACTAGGAGGGGCTCGGCCTCGCcag
cttccgtagccccgccccgccccgctccgcttcggacctctgctgggtccccagggact
cggctgtgcgcgtgagagtaaagccagatcgtaagagaaaagttcttcccccgtttctt
cttctccggacgtcgcccagccttctgcctctcggctgccgagttcccacaggctctgg
gagactgaggctgccagggtcagactaaagagaggtctcagagagtttaattcaacact
tcttggctactaagtcttagaagtctgatggtgtgctctctctgctgagttggggagcg
tgaatggaggctatgtcaccgaagctgatagagctcagtctctgttgcagatgctcccg
acccttttgcattgggccagttccccagctctgagactgggtccaggctcaggaagtgg
cctatgtgtcaaggtggattc

FIGURE 23w

45080
GAATGGTCATTTTTGATGTTTTGTTGTTGTTGCTATTTTCGTTGTTGAGGATAACTATA
ATTTTTTGTGCCAAAAATGTGGCAAACCTTTCTATGGGGAAAACGATAGAAATGGCACT
TAACCCTAACCCATTGGACATAATCTATTATCTGTTTTTACTAAAATCCACTGAACCTG
TAGAAATCTTAGATTAATCAGAAACACACTCTTTTCTTGTGCTTCTCAATAAATAATTG
AATTGTTTTTGCCCAGGAATTACCCCTGAGCAACTAAAATGTTTACCTTCCTGCAGTTA
TAAAAATCTCGGTGGGGGTTGTTTTTCAGCTCCTTTAACTCGTCCATCTCGTTAAGCAT
CTGATGGACCTGGAACTTGGAGGAGAGGAACTTCAGGCGCCGGTGGGTATAGGTCTTAC
TGTGAAAAATAAAATCACATAATTCCAAAAAGTTTCAGGCATTCAAGAAAAACAGTCAC
AATTTCAAAACTATCAGGACCTTTATCATTCATAGGAAATAATTGTTGGAACAAACCTT
TTAGTTTACTCTGCAGTTAATCCCACTGA[ga]AAGTAGTGGGCTCCAAAGGCTTAATC
TTTTCAATAATGTTGGACATAAGAATGAGGGAGAACTTGGAAAGGTATCTTAAAACTCA
ATGGAGAGAGTGTTATTCAAAGTTTGGGGTCAGCAGATTCGAGTGTGAATCCTGGCTCA
GCCAGCTGTGTCACTTTAGGCAAGTTACTTAAGTCATCAAAGTCTCAGCTCATAAAACT
GGAATTATGAAAATAACCACCTCACAGTGAAAAGTGTAAGCAATAAAAGGAACAATGTG
CATGAAGGGCTTAATACAGTGTTTGAACATAGTAAGCATTTAGTAAATACTTAGTCTCA
CTATCAGTAGAAGTAGTACTAGTTGTTGTTTAGGTCTTGTAGTACTAGTTGTTGTTGTT
TAGGTCTCACTAAACACTTACACAGGTCCTTGAGCAATTAAAGCAAGTAAAAAATTCAT
ATCGTCTAAGAAGGTGTCCAG

gaatggtcatttttgatgttttgttgttgttgctattttcgttgttgaggataactata
attttttgtgccaaaaatgtggcaaacctttctatggggaaaacgatagaaatggcact
taaccctaacccattggacataatctattatctgtttttactaaaatccactgaacctg
tagaaatcttagattaatcagaaacacactcttttcttgtgcttctcaataaataattg
aattgtttttgcccaggaattacccctgagcaactaaaatgtttaccttcctgcagtta
taaaaatctcggtgggggttgtttttcagctcctttaactcgtccatctcgttaagcat
ctgatggacctggaacttggaggagaggaacttcaggcgccggtgggtataggtcttac
tgtgaaaaataaaatcacataattccaaaaagtttcaggcattcaagaaaaacagtcac
aatttcaaaactatcaggacctttatcattcataggaaataattgttGGAACAAACCTT
TTAGTTTACTCTGCAGTTAATCCCACTGA[ga]AAGTAGTGGGCTCCAAAGGCTTAATC
TTTTCAATAATGTTGgacataagaatgagggagaacttggaaaggtatcttaaaactca
atggagagagtgttattcaaagtttggggtcagcagattcgagtgtgaatcctggctca
gccagctgtgtcactttaggcaagttacttaagtcatcaaagtctcagctcataaaact
ggaattatgaaaataaccacctcacagtgaaaagtgtaagcaataaaaggaacaatgtg
catgaagggcttaatacagtgtttgaacatagtaagcatttagtaaatacttagtctca
ctatcagtagaagtagtactagttgttgtttaggtcttgtagtactagttgttgttgtt
taggtctcactaaacacttacacaggtccttgagcaattaaagcaagtaaaaaattcat
atcgtctaagaaggtgtccag

FIGURE 23x

47848
GAAAGCTGAGAAAGAGGCACACCAAGACTAAGGGAAAGAGGCCGGGAAGGGTAAAAAGG
TGAAATGAAAAGAGGTTGGTGAATGACTAAGAACGGTTGGATAGGACAAATAAGTTCCA
ATGTTCGATAGCAGACGAGGGTGACTACAGTTAGCAATATATTGTATATTTCAAAGTAG
CTAGAAGACTTAAAATGTTATCAACACATAGAAATGAAATATACCTAAGGTGATGTATC
CTTCAAATACCCGGACTTGATCATTACACATTCCGGGCATGTAAAAAACGCTTCCATGT
ACCCCATTTCATAAATATGTAAAATATTATGTATCATTAAAAGAAAGAACAAAAAAGAC
AGGGAAAATGCATATGCTGTGCTCCACTCAGCCAACAAACTTCTGCTCTAAGCAGGGAT
ATTGATTCCAAAGGCTAGCTTGCGTTTCTTAAAAATAATTAAAAACAACAACATGTCAT
TTATTTCAGAGCTGGAGGCTAGAAATAAATTACTCAAATCTCGCAACTATGTAAACTAT
GAAAATGAAACAAGCTAGTTACCTTTTAT[tc]GTTCAGTTTAAAAAAGTTCTTCTTCT
TTGCTCCTCCATTGCGGTCCCCTTCAAGATCCATTCCGACCTGAAGAGAAACCGCAGCT
CATTAGCCAAATGCATGAGCCTCAGGCGCGCTGGAGGTGAGACTAACCTCTAGTCCCCC
GTCGAAGCCAGAGAGCAGTAAGAGGGAGCGCCCGCCGTTGATGCCCCAGCTGCTCTGGC
CGCGATGGGCACTGCAGGGGCTTTCCTGTGCGCGGGGTCTCCAGCATCTCCACGAAGGC
AGAGTTGGGGGTCTGGCAGCGCGTTCTGGACTTTGCCCGCCGCCAGTGCGATTCTCCCT
CCCGGTTCCAGTCGCCGCGGACGATGCTTCCTCCCACCCACCGCCCGCGGGCTCAGAGA
GCAGGTCCCCGCACCGCGCGGGCTGTGCGCGCTCCGGGCAACATGGTCCAGTGCCACTA
CGGTTTGGGCGCTGCTCCAGG

gaaagctgagaaagaggcacaccaagactaagggaaagaggccgggaagggtaaaaagg
tgaaatgaaaagaggttggtgaatgactaagaacggttggataggacaaataagttcca
atgttcgatagcagacgagggtgactacagttagcaatatattgtatatttcaaagtag
ctagaagacttaaaatgttatcaacacatagaaatgaaatatacctaaggtgatgtatc
cttcaaatacccggacttgatcattacacattccgggcatgtaaaaaacgcttccatgt
accccatttcataaatatgtaaaatattatgtatcattaaaagaaagaacaaaaaagac
agggaaaatgcatatgctgtgctccactcagccaacaaacttctgctctaagcagggat
attgattccaaaggctagcttgcgtttcttaaaaataattaaaaacaacaacatgtcat
ttatttcagagctggaggctagaaataaattactcaaatctcGCAACTATGTAAACTAT
GAAAATGAAACAAGCTAGTTACCTTTTAT[tc]GTTCAGTTTAAAAAAGTTCTTCTTCT
TTGCTCCTCCATTGCGGTCCccttcaagatccattccgacctgaagagaaaccgcagct
cattagccaaatgcatgagcctcaggcgcgctggaggtgagactaacctctagtccccc
gtcgaagccagagagcagtaagagggagcgcccgccgttgatgccccagctgctctggc
cgcgatgggcactgcaggggctttcctgtgcgcggggtctccagcatctccacgaaggc
agagttgggggtctggcagcgcgttctggactttgcccgccgccagtgcgattctccct
cccggttccagtcgccgcggacgatgcttcctcccacccaccgcccgcgggctcagaga
gcaggtccccgcaccgcgcgggctgtgcgcgctccgggcaacatggtccagtgccacta
cggtttgggcgctgctccagg

FIGURE 23y

47850
GTGAGTTTTGAGGCTTGGGAGAGAGCTGCAAGGAGGAAGAAGGAAGAGAAATAGGGGAG
AGACATGGGGAGAGACAGTCATGCCTACTTCCTCAGCAGGCCAGAAGCAGCATGTGCAG
GTGGGGACCCAGACTCTGTACTTGGACTTAAAGTGAAAGGCTTTCCAGATATTGTACTT
ACCCCTAAGGCTGACAAAGGTGGAGCCTCAAGCCTATAGCTTTGGATCAAGACAATTGT
TCCAGTTCTCCTATCCCAGAAATGTTCCTCTCTCCTAAACCTGAAGTGGTCGAACACTT
TCATCCCTTCCTCACAAGGAGGGTCAGGTGATCAGGTAAAGGTAACAACTAACCCAAAC
AGGAAGTGTGGCCAGATGCTTGTATACAGGTAAGGGTGTGATTTGGTTGCTAATTTCTC
TTCACTTCTGGGAGACCAGCCCCTTATAAATCAAACTATAGGCCAGAGAGGCTGCCACA
TGCTCCCAGGCTGTTTATTTGAAGAGAGACTTACATTAGGCAGTGACTCGATGAAGGCA
TGTATGTTGGCCTCCTTTGCTGCCCTCAC[ga]ATCTCTTCCTGTGACACCACCCGGCT
GTTGTCTCCATAGGCAATGTTCTCAGCAATGCTGCAGTCAAACAGGATGGGCTCCTGGG
ACACGATGCCCAGGTGTGCTCGGAGCCACTGAACATTCAGTCGCTTTATTTCTTTGCCA
TCAAGCAGCTGAAAACAAGAGTTCACAGATCAACTTCAGGACCAGCACACTTTGAATGT
AGCACAATTAACATCATTATTTCTTACACTGAAACTGCCAAGTTACTGTGAGATTAAGG
AAAAGTTTGTGTGATTAAAATTTGGATAGTGAAGGTTAACCCAACAAGGTCATAATTGT
ATGCCTTGAGGAACTGTCATGTTTCCTGTGTTTCAACCATGGTTTCTGATGTATGCATG
TGGTAGGCAGAATAATGTTCCCTCTCCCACAAGACATCTGTGTCCTAATCCCTGGATCC
TGTGAATGTGTTATGTTACAT

gtgagttttgaggcttgggagagagctgcaaggaggaagaaggaagagaaataggggag
agacatggggagagacagtcatgcctacttcctcagcaggccagaagcagcatgtgcag
gtggggacccagactctgtacttggacttaaagtgaaaggctttccagatattgtactt
acccctaaggctgacaaaggtggagcctcaagcctatagctttggatcaagacaattgt
tccagttctcctatcccagaaatgttcctctctcctaaacctgaagtggtcgaacactt
tcatcccttcctcacaaggagggtcaggtgatcaggtaaaggtaacaactaacccaaac
aggaagtgtggccagatgcttgtatacaggtaagggtgtgatttggttgctaatttctc
ttcacttctgggagaccagccccttataaatcaaactataggccagagaggctgccaca
tgctcccaggctgtttatttgaagagagacttacattaggcagtgactcgatgaaggcA
TGTATGTTGGCCTCCTTTGCTGCCCTCAC[ga]ATCTCTTCCTGTGACACCACCCGGCT
GTTGtctccataggcaatgttctcagcaatgctgcagtcaaacaggatgggctcctggg
acacgatgcccaggtgtgctcggagccactgaacattcagtcgctttatttctttgcca
tcaagcagctgaaaacaagagttcacagatcaacttcaggaccagcacactttgaatgt
agcacaattaacatcattatttcttacactgaaactgccaagttactgtgagattaagg
aaaagtttgtgtgattaaaatttggatagtgaaggttaacccaacaaggtcataattgt
atgccttgaggaactgtcatgtttcctgtgtttcaaccatggtttctgatgtatgcatg
tggtaggcagaataatgttccctctcccacaagacatctgtgtcctaatccctggatcc
tgtgaatgtgttatgttacat

FIGURE 23z

47852
GCCTTGCCTTCCCCCAGGCAGGTTTGAGAGGTCTGGGTCTCAACTGACTGGGGCAGCAG
GACCTCATCCCCTCCCTGCCCTACACCCAGCCTGCCCCAGCCCTGCAGTCTGTTGTTCC
TTAGTCAGGGAGGAGCCCAAAAGTGTGACCAAACCAAGGAACACTCAACTTCTGGCTT
CCTCCCTCTTTGGGTAGCCCTCAAGCCACTGGACTTTGAAGTCAGCAGGTAATTCTCCA
AATGGAAGAACTTTTTTTTTTTTTTTAAAAGCAGAGCCAAGGAAGCCACATTTTGAGT
GATGTGGTTTTTGAAGAAAAAAGAAAAAGAGATCCCAGATAAAAATGATCTTATGTGAA
GGGAGTAAATGGATGCACAGAAACAGCAGCAGCTCCCGAGCCACCTGGTGGAGCACAGG
GGCCCTCCCTGGCCTCCCCCAACACTGGGGCTGGGGTCTGGGGGCTGCCCAGCAGGGTG
ATGTGGCTCCCTTGGGCCTGAGAGCACCCTGGAGGGAGTTGACCCTGGGGGGCAATGTT
CCCAGGACGCAGTACCTGATATCCAAGTC [ga] GTCGCTGTCTCCCGCTCTGGGCTGCA
GCAGGGGAGGAAAGGCATACTGAGCTCTCATGGGAGTGAACCATATCCTCCAGGAAGAT
CCTGAGCTCCCTCCAACCCAACATGAGCATGCCTTTACAATCCCCTGGACCCAGTCTGT
AGCCACAAATGCTGCATAGAGAGGTGTGGAGAGTGGGGTGTGCCCATCTTGGGGAAGCC
TCTGCTGCCTGACCACGTGGGTGTGTGAGGAGGGCCCTGGAGGACCCAGTTAAGAGGGA
GAATGGGGAGAGGTGCCATTGGTGCAGGCTCTGGGGGGAAAACTTGTCAGATCAGGAGT
ATGAAGCCCGCAATGTGGCTCCTCCAGACCCAGCCTCTGCATTCAGGTTGGAATGAATA
GGCTGAGGTCTGAGGCTGATACAGCTGCACAAACAGCTGGGGCAAGGAGTGCTCTGGAC
AGAGCCAGGCCAGGCCAGGCA

gccttgccttcccccaggcaggtttgagaggtctgggtctcaactgactggggcagcag
gacctcatcccctccctgccctacacccagcctgccccagccctgcagtctgttgttcc
ttagtcagggaggagcccaaaagtgtgaccaaaccaaggaacactcaacttctggctt
cctccctctttgggtagccctcaagccactggactttgaagtcagcaggtaattctcca
aatggaagaactttttttttttttttaaaagcagagccaaggaagccacattttgagt
gatgtggtttttgaagaaaaaagaaaaagagatcccagataaaaatgatcttatgtgaa
gggagtaaatggatgcacagaaacagcagcagctcccgagccacctggtggagcacagg
ggccctccctggcctcccccaacactggggctggggtctgggggctgcccagcagggtg
atgtggctcccttgggcctgagagcaccctggagggagttgaccctggggggcaatgtt
cccagGACGCAGTACCTGATATCCAAGTC [ga] GTCGCTGTCTCCCGCTCTGGGCTGca
gcaggggaggaaaggcatactgagctctcatgggagtgaaccatatcctccaggaagat
cctgagctccctccaacccaacatgagcatgcctttacaatcccctggacccagtctgt
agccacaaatgctgcatagagaggtgtggagagtggggtgtgcccatcttggggaagcc
tctgctgcctgaccacgtgggtgtgtgaggagggccctggaggacccagttaagaggga
gaatggggagaggtgccattggtgcaggctctggggggaaaacttgtcagatcaggagt
atgaagcccgcaatgtggctcctccagacccagcctctgcattcaggttggaatgaata
ggctgaggtctgaggctgatacagctgcacaaacagctggggcaaggagtgctctggac
agagccaggccaggccaggca

FIGURE 23aa

47854
TGTCAGGCAAGATTCAAATCAAAATAATTAATTTTAAATGACATGCATACTTTTTGGAG
AGAAAAGTTTGGGTTACAATTAGCCAATCTGTTAAAACTCAAAGAAATCTAATCCAAAC
GTAATACACATGTCTGTACCATTTTTTTAGCCTATTCTCTCTTCAGACTTATACTTAA
TCACAAATAACATTCTTCTTTCTATTAATTAATTCCAAAAACTGGCTCACAGCCATATA
TGACAGTCATTTATTGCTACTAGGGACATAAAATTTCTAAATAATCAGAAATCCACGTT
GTCATTTATGAATATTCTCTCTCCTTGCAAACCAAAAAAATCATCTTTAACCTTACCTG
ATAGATTTTGGCATCCCTCATTAGTTTTCTACAGGATATTCTGTATTAAATCCATTGC
CTCCAAGTATCTGCACAGCATCAGTAGCTAACTGATTTGCAATATCTCCAGCAAATGCC
TTTGCAATAGAAGCATAATAGGTATTTCGACGACCAGAATCAACCTCCCAAGCTGCTCT
CTGGTAACTCATTCTAGCTAGTTCAACTT[tc]CATTGCCATTTCAGCCAGCATAAATG
ATATTGCTTGGTGCTAGAATTAAAAAGAAAAAAATTAAAGGATATTTATTGAGAAAACT
TAAAAGTTTTTTCCTGGGGCTTTTTCATTTTTATAGTGACGGGGTCTTGCTATGTTGCC
CAGGCTGGTCTGCAACTCCTGGCCTCAAGCAATCCTCCTACTTAGGCCTCTCAAAGTGC
TGAGATTACAGGCGTGAGCCACTGTGCCTGACCTTTTTATTTTTTAAACTTTTCATTAA
CGAATTTTAGGTTTATAGAAGTTACACCCAGCTTCCTCTAATGTTAACATATTACCAAA
CCATAGTGCCATGATCGAGAACAGGACATTAACACTGGTATAGTATTAACAACTAAACT
ATAAGCCTTACTCAAATCTGGTCAAGTTTTCTACTAATGTTCTTTTTCCACCATTATAC
GTTGAATTTAGTTATTTCTTC

tgtcaggcaagattcaaatcaaaataattaattttaaatgacatgcatactttttggag
agaaaagtttgggttacaattagccaatctgttaaaactcaaagaaatctaatccaaac
gtaatacacatgtctgtaccatttttttagcctattctctcttcagacttatacttaa
tcacaaataacattcttctttctattaattaattccaaaaactggctcacagccatata
tgacagtcatttattgctactagggacataaaatttctaaataatcagaaatccacgtt
gtcatttatgaatattctctctccttgcaaaccaaaaaaatcatctttaaccttacctg
atagattttggcatccctcattagttttctacaggatattctgtattaaatccattgc
ctccaagtatctgcacagcatcagtagctaactgatttgcaatatctccagcaaatgcc
tttgcaatagaagcataataggtatttcgacgaccagaatcaacctcCCAAGCTGCTCT
CTGGTAACTCATTCTAGCTAGTTCAACTT[tc]CATTGCCATTTCAGCCAGCATAAATG
ATATTGCTTGGTGCTagaattaaaaagaaaaaaattaaaggatatttattgagaaaact
taaaagttttttcctggggctttttcatttttatagtgacggggtcttgctatgttgcc
caggctggtctgcaactcctggcctcaagcaatcctcctacttaggcctctcaaagtgc
tgagattacaggcgtgagccactgtgcctgacctttttattttttaaacttttcattaa
cgaattttaggtttatagaagttacacccagcttcctctaatgttaacatattaccaaa
ccatagtgccatgatcgagaacaggacattaacactggtatagtattaacaactaaact
ataagccttactcaaatctggtcaagttttctactaatgttctttttccaccattatac
gttgaatttagttatttcttc

FIGURE 23bb

47855
AGGTAGCGGCCACAGAAGAGCCAAAAGCTCCCGGGTTGGCTGGTAAGGACACCACCTCC
AGCTTTAGCCCTCTGGGGCCAGCCAGGGTAGCCGGGAAGCAGTGGTGGCCCGCCCTCCA
GGGAGCAGTTGGGCCCCGCCCGGGCCAGCCCCAGGAGAAGGAGGGCGAGGGGAGGGGAG
GGAAAGGGGAGGAGTGCCTCGCCCCTTCGCGGCTGCCGGCGTGCCATTGGCCGAAAGTT
CCCGTACGTCACGGCGAGGGCAGTTCCCCTAAAGTCCTGTGCACATAACGGGCAGAACG
CACTGCGAAGCGGCTTCTTCAGAGCACGGGCTGGAACTGGCAGGCACCGCGAGCCCCTA
GCACCCGACAAGCTGAGTGTGCAGGACGAGTCCCCACCACACCCACACCACAGCCGCTG
AATGAGGCTTCCAGGCGTCCGCTCGCGGCCCGCAGAGCCCCGCCGTGGGTCCGCCCGCT
GAGGCGCCCCCAGCCAGTGCGCTCACCTGCCAGACTGCGCGCCATGGGGCAACCCGGGA
ACGGCAGCGCCTTCTTGCTGGCACCCAAT[ag]GAAGCCATGCGCCGGACCACGACGTC
ACGCAGGAAAGGGACGAGGTGTGGGTGGTGGGCATGGGCATCGTCATGTCTCTCATCGT
CCTGGCCATCGTGTTTGGCAATGTGCTGGTCATCACAGCCATTGCCAAGTTCGAGCGTC
TGCAGACGGTCACCAACTACTTCATCACTTCACTGGCCTGTGCTGATCTGGTCATGGGC
CTGGCAGTGGTGCCCTTTGGGGCCGCCCATATTCTTATGAAAATGTGGACTTTTGGCAA
CTTCTGGTGCGAGTTTTGGACTTCCATTGATGTGCTGTGCGTCACGGCCAGCATTGAGA
CCCTGTGCGTGATCGCAGTGGATCGCTACTTTGCCATTACTTCACCTTTCAAGTACCAG
AGCCTGCTGACCAAGAATAAGGCCCGGGTGATCATTCTGATGGTGTGGATTGTGTCAGG
CCTTACCTCCTTCTTGCCCAT

aggtagcggccacagaagagccaaaagctcccgggttggctggtaaggacaccacctcc
agctttagccctctggggccagccagggtagccgggaagcagtggtggcccgccctcca
gggagcagttgggccccgcccgggccagccccaggagaaggagggcgaggggaggggag
ggaaaggggaggagtgcctcgccccttcgcggctgccggcgtgccattggccgaaagtt
cccgtacgtcacggcgagggcagttcccctaaagtcctgtgcacataacgggcagaacg
cactgcgaagcggcttcttcagagcacgggctggaactggcaggcaccgcgagccccta
gcacccgacaagctgagtgtgcaggacgagtccccaccacacccacaccacagccgctg
aatgaggcttccaggcgtccgctcgcggcccgcagagccccgccgtgggtccgcccgct
gaggcgcccccagccagtgcgctcacctgccagactgcgcgccatggggcaacccggga
acggcaGCGCCTTCTTGCTGGCACCCAAT[ag]GAAGCCATGCGCCGGACCACGACgtc
acgcaggaaagggacgaggtgtgggtggtgggcatgggcatcgtcatgtctctcatcgt
cctggccatcgtgtttggcaatgtgctggtcatcacagccattgccaagttcgagcgtc
tgcagacggtcaccaactacttcatcacttcactggcctgtgctgatctggtcatgggc
ctggcagtggtgccctttggggccgcccatattcttatgaaaatgtggacttttggcaa
cttctggtgcgagttttggacttccattgatgtgctgtgcgtcacggccagcattgaga
ccctgtgcgtgatcgcagtggatcgctactttgccattacttcacctttcaagtaccag
agcctgctgaccaagaataaggcccgggtgatcattctgatggtgtggattgtgtcagg
ccttacctccttcttgcccat

FIGURE 23cc

47857
CAGAGCCCCGCCGTGGGTCCGCCCGCTGAGGCGCCCCCAGCCAGTGCGCTCACCTGCCA
GACTGCGCGCCATGGGGCAACCCGGGAACGGCAGCGCCTTCTTGCTGGCACCCAATaGA
AGCCATGCGCCGGACCACGACGTCACGCAGGAAAGGGACGAGGTGTGGGTGGTGGGCAT
GGGCATCGTCATGTCTCTCATCGTCCTGGCCATCGTGTTTGGCAATGTGCTGGTCATCA
CAGCCATTGCCAAGTTCGAGCGTCTGCAGACGGTCACCAACTACTTCATCACTTCACTG
GCCTGTGCTGATCTGGTCATGGGCCTGGCAGTGGTGCCCTTTGGGGCCGCCCATATTCT
TATGAAAATGTGGACTTTTGGCAACTTCTGGTGCGAGTTTTGGACTTCCATTGATGTGC
TGTGCGTCACGGCCAGCATTGAGACCCTGTGCGTGATCGCAGTGGATCGCTACTTTGCC
ATTACTTCACCTTTCAAGTACCAGAGCCTGCTGACCAAGAATAAGGCCCGGGTGATCAT
TCTGATGGTGTGGATTGTGTCAGGCCTTA[ct]CTCCTTCTTGCCCATTCAGATGCACT
GGTACCGGGCCACCCACCAGGAAGCCATCAACTGCTATGCCAATGAGACCTGCTGTGAC
TTCTTCACGAACCAAGCCTATGCCATTGCCTCTTCCATCGTGTCCTTCTACGTTCCCCT
GGTGATCATGGTCTTCGTCTACTCCAGGGTCTTTCAGGAGGCCAAAAGGCAGCTCCAGA
AGATTGACAAATCTGAGGGCCGCTTCCATGTCCAGAACCTTAGCCAGGTGGAGCAGGAT
GGGCGGACGGGGCATGGACTCCGCAGATCTTCCAAGTTCTGCTTGAAGGAGCACAAAGC
CCTCAAGACGTTAGGCATCATCATGGGCACTTTCACCCTCTGCTGGCTGCCCTTCTTCA
TCGTTAACATTGTGCATGTGATCCAGGATAACCTCATCCGTAAGGAAGTTTACATCCTC
CTAAATTGGATAGGCTATGTC

cagagccccgccgtgggtccgcccgctgaggcgcccccagccagtgcgctcacctgcca
gactgcgcgccatggggcaacccgggaacggcagcgccttcttgctggcacccaataga
agccatgcgccggaccacgacgtcacgcaggaaagggacgaggtgtgggtggtgggcat
gggcatcgtcatgtctctcatcgtcctggccatcgtgtttggcaatgtgctggtcatca
cagccattgccaagttcgagcgtctgcagacggtcaccaactacttcatcacttcactg
gcctgtgctgatctggtcatgggcctggcagtggtgccctttggggccgcccatattct
tatgaaaatgtggacttttggcaacttctggtgcgagttttggacttccattgatgtgc
tgtgcgtcacggccagcattgagaccctgtgcgtgatcgcagtggatcgctactttgcc
attacttcacctttcaagtaccagagcctgctgaccaagaataaggcccgggtgaTCAT
TCTGATGGTGTGGATTGTGTCAGGCCTTA[ct]CTCCTTCTTGCCCATTCAGATGCACT
GGTACCGggccacccaccaggaagccatcaactgctatgccaatgagacctgctgtgac
ttcttcacgaaccaagcctatgccattgcctcttccatcgtgtccttctacgttcccct
ggtgatcatggtcttcgtctactccagggtctttcaggaggccaaaaggcagctccaga
agattgacaaatctgagggccgcttccatgtccagaaccttagccaggtggagcaggat
gggcggacggggcatggactccgcagatcttccaagttctgcttgaaggagcacaaagc
cctcaagacgttaggcatcatcatgggcactttcaccctctgctggctgcccttcttca
tcgttaacattgtgcatgtgatccaggataacctcatccgtaaggaagtttacatcctc
ctaaattggataggctatgtc

FIGURE 23dd

47867
CCACTCCGGAGCACCTGGCTCTGCCCTCAGGAACTCCCTGAGCTTTGCACACAGGGCCG
AGACACCTGGATTTCTCTGGTTCCCTGAGTGGGGCCAGCTTGGAAGAATTTCCCAAAGC
CTATTAGAGCAACGGCTGCCTCCTGCCTGCCTCCTTGGGCTGGGCAGGGCTGAGGGCGG
AGGGAGAGAGAGAGAGAGGGAGGGGGAGAGGAGGAAGGAAAAAGTTGGCAGGCCGACAG
CACAGCCGTGTCTGCATCCATCCAGAGGAGGTCTGTGTGGTGTGGGGCGGGCCAGGAGC
GAAGAGAGGCCTTCCTCCCTTTGTGCTCCCCCCGCCCCCCGGCCCTATAAATAGGCCCA
GCCCAGGCTGTGGCTCAGCTCTCAGAGGGAATTGAGCACCCGGCAGCGGTCTCAGGCCA
AGCCCCCTGCCAGCATGGCCAGCGAGTTCAAGAAGAAGCTCTTCTGGAGGGCAGTGGTG
GCCGAGTTCCTGGCCACGACCCTCTTTGTCTTCATCAGCATCGGTTCTGCCCTGGGCTT
CAAATACCCGGTGGGGAACAACCAGACGG[ct]GGTCCAGGACAACGTGAAGGTGTCGC
TGGCCTTCGGGCTGAGCATCGCCACGCTGGCGCAGAGTGTGGGCCACATCAGCGGCGCC
CACCTCAACCCGGCTGTCACACTGGGGCTGCTGCTCAGCTGCCAGATCAGCATCTTCCG
TGCCCTCATGTACATCATCGCCCAGTGCGTGGGGGCCATCGTCGCCACCGCCATCCTCT
CAGGCATCACCTCCTCCCTGACTGGGAACTCGCTTGGCCGCAATGACGTGAGTGGGGTG
TCCCTGGGCTTGGGGGGGTTCTAGAATGATGCTGAAAGGCACTGGTTCCATCCTCTGCC
CATTGTGCAGATGGGGACACTGAGGAACGGAGAGGACAAGAGGTTGCTGGAGGTCACGT
AGAGAGCTGGGGGGAAGAGCTGGGGCTGGAACTCAGCTATGCATGCCTCCCAAAGCCTG
TTTTCTGCCAGGCACTGTGGG

ccactccggagcacctggctctgccctcaggaactccctgagctttgcacacagggccg
agacacctggatttctctggttccctgagtggggccagcttggaagaatttcccaaagc
ctattagagcaacggctgcctcctgcctgcctccttgggctgggcagggctgagggcgg
agggagagagagagagagggagggggagaggaggaaggaaaaagttggcaggccgacag
cacagccgtgtctgcatccatccagaggaggtctgtgtggtgtggggcgggccaggagc
gaagagaggccttcctccctttgtgctccccccgccccccggccctataaataggccca
gcccaggctgtggctcagctctcagagggaattgagcacccggcagcggtctcaggcca
agccccctgccagcatggccagcgagttcaagaagaagctcttctggagggcagtggtg
gccgagttcctggccacgaccctctttgtcttcatcagcatcggttctgccctgggctt
cAAATACCCGGTGGGGAACAACCAGACGG[ct]GGTCCAGGACAACGTGAAGGTGTCGC
TGgccttcgggctgagcatcgccacgctggcgcagagtgtgggccacatcagcggcgcc
cacctcaacccggctgtcacactggggctgctgctcagctgccagatcagcatcttccg
tgccctcatgtacatcatcgcccagtgcgtgggggccatcgtcgccaccgccatcctct
caggcatcacctcctccctgactgggaactcgcttggccgcaatgacgtgagtggggtg
tccctgggcttgggggggttctagaatgatgctgaaaggcactggttccatcctctgcc
cattgtgcagatggggacactgaggaacggagaggacaagaggttgctggaggtcacgt
agagagctggggggaagagctggggctggaactcagctatgcatgcctcccaaagcctg
ttttctgccaggcactgtggg

FIGURE 23ee

47868
CTCCTCACCAGTCCTCACCACCTCTCTCCCCTGCAGCTGGCTGATGGTGTGAACTCGGG
CCAGGGCCTGGGCATCGAGATCATCGGGACCCTCCAGCTGGTGCTATGCGTGCTGGCTA
CTACCGACCGGAGGCGCCGTGACCTTGGTGGCTCAGCCCCCCTTGCCATCGGCCTCTCT
GTAGCCCTTGGACACCTCCTGGCTGTGAGTCAGGGGCCCTCCCAGATGGAGGTGGGGGA
AGGGAGGGCGGGGGCTGGTGGGGTGCCCTGCCATGGGCAGCCAGTGGGACTCCCGACAG
GGCTCTTGCCATTGGGTGGAGGATGGCGGGTCAGCGCTGGGGGCTGGGGGCAGGGTCCT
GCCCTGGAGAGGAGCACAGGGACCTCCTGCCCAGCTTGGGGTCAGCACTCCTCTTTCCC
TGGGTCTCATTGTCCCCCACCCTGATTGTTCTCTTTCTCCCTCCAACCTCTCCCTCCTC
TCACTCTCTCTTCACCTATGACTCTCTGCCTTCGCCCCTCCCTCTGTTTCTTTCCCTCA
CAGATTGACTACACTGGCTGTGGGATTAA[ac]CCTGCTCGGTCCTTTGGCTCCGCGGT
GATCACACACAACTTCAGCAACCACTGGGTAGGAGACCCACGGGGGGTGGGGTGGGAAG
CTTTGGTGTCCCATGGTAAGCCTGACCCCACCCTCACAGTGTCCCTTCCTGTTCTGGAG
GCTCTGGGAGACAGCCAGAGGACAGGAAATCAGGAAACTGAGGCCTGCCATGTAGAGGC
AGGCTGGGGGTCACACTGCCAGCACTTTCAGGCCTAGTCTCTGCCCTCCCAGCTCGGCC
CTGCCCCATGCTGCCTGGCCTCCAGGTCTTCCCAGCTGCGTGGTTAAAAGTGGGGCTCC
AAATCCTGGCTCAGCCACTTTCGGGTTTAGCATGACCTTGCGCAGTGTGCTTGAGCTTT
GGTTTCCTGAGCTGCGGAGGGGGATATGGTGGTGCCCACCTCTCAGGGTGGCCGAGAAG
AGGAAAGGGCTCACTCCCCAT

ctcctcaccagtcctcaccacctctctcccctgcagctggctgatggtgtgaactcggg
ccagggcctgggcatcgagatcatcgggaccctccagctggtgctatgcgtgctggcta
ctaccgaccggaggcgccgtgaccttggtggctcagccccccttgccatcggcctctct
gtagcccttggacacctcctggctgtgagtcaggggccctcccagatggaggtgggggga
agggagggcgggggctggtggggtgccctgccatgggcagccagtgggactcccgacag
ggctcttgccattgggtggaggatggcgggtcagcgctgggggctgggggcagggtcct
gccctggagaggagcacagggacctcctgcccagcttggggtcagcactcctctttccc
tgggtctcattgtcccccaccctgattgttctctttctccctccaacctctccctcctc
tcactctctcttcacctatgactctctgccttcgcccctccctctgtttctttccctca
cagatTGACTACACTGGCTGTGGGATTAA[ac]CCTGCTCGGTCCTTTGGCTCCGCGgt
gatcacacacaacttcagcaaccactgggtaggagacccacggggggtggggtgggaag
ctttggtgtcccatggtaagcctgaccccaccctcacagtgtcccttcctgttctggag
gctctgggagacagccagaggacaggaaatcaggaaactgaggcctgccatgtagaggc
aggctgggggtcacactgccagcactttcaggcctagtctctgccctcccagctcggcc
ctgccccatgctgcctggcctccaggtcttcccagctgcgtggttaaaagtggggctcc
aaatcctggctcagccactttcgggtttagcatgaccttgcgcagtgtgcttgagcttt
ggtttcctgagctgcggagggggatatggtggtgcccacctctcagggtggccgagaag
aggaaagggctcactccccat

FIGURE 23ff

47877
TTTGACTCCCTGTACCTTTAAGAGGGACCCTTAAATTTAAAAATCTATTGTATTTTTTT
TTTAGTAGGGGTAGGGAATATTTAGGGAATTTGGAAGGGGTTATATAGTTCTTTAAGAA
TCAAATAGCACATCTTCCTGAAAATAGCACGTAGACAAAGTTTTTTTGGAGATAACCTT
AGGAATATCGTAACTCTCTGATGCCACCTCCATATGTGATCCTATGTTGATTATAAGAT
TTTGATCAGTGGCTTTCAGACTTTTTTGACTGCAACCTAGAATAAAAGATTCATTTACA
TTGTGACCTAGAACACACACACACACACACACTCTCTCTCCGCCACTCTCCTGCACACA
GAAATCATTGATGCTTACAACAATTCTTACTCTTACTATGGGTGATTTACTTTGATATG
CTCTGTTTTTTTTTTCATTTACAAAACTGTGGATTAATTTTTTTTGACATGCTAAATTG
ATCTCAGTAATAGATTGTATTTATTCTTCCTTAGATTCTTCTTTGGAGCAGAATAAAAG
ATCTGGCCCATCAGTTCACACAGGTCCAG[ct]GGGACATGTTCACCCTGGAGGACACG
CTGCTAGGCTACCTTGCTGATGACCTCACATGGTGTGGTGAATTCAACACTTCCAGTGA
GGCTCTGGGCCCTGTGGGATTGCCCAGGGATGTGGAGGGTGAACAGAGTGACTTCTGCT
GGAGGCCCTGAATGATTAGTGTGGAGGACAGAGCCACAGGCACCCATCCTGATGCCATC
TATACTTATATTAGTCCATTTGTGTTGCTATTAAGGAATACCTGAGGCTGCGTAATTTA
TAAAGAAAAGAGGTTTATTTGACTCACAGTTACGCAGGCTGTACAAGAAGTAGGGTACC
AGCATCCACTTCGGGTGAAGGCCTGAGGCTGTTTCCACTCATGGAGAAGGGGAAGGGGA
GCTGGCATTTACAGAGATCACATGGTGAGGGAGGAAAGCAAGGAGAGGTCAGGGGAGGT
GCCAGGCTGTTTGTAATGACC

tttgactccctgtacctttaagagggacccttaaatttaaaaatctattgtattttttt
tttagtaggggtagggaatatttagggaatttggaaggggttatatagttctttaagaa
tcaaatagcacatcttcctgaaaatagcacgtagacaaagtttttttggagataacctt
aggaatatcgtaactctctgatgccacctccatatgtgatcctatgttgattataagat
tttgatcagtggctttcagacttttttgactgcaacctagaataaaagattcatttaca
ttgtgacctagaacacacacacacacacacactctctctccgccactctcctgcacaca
gaaatcattgatgcttacaacaattcttactcttactatgggtgatttactttgatatg
ctctgttttttttttcatttacaaaactgtggattaattttttttgacatgctaaattg
atctcagtaatagattgtatttattcttccttagattcttctttggagcagaataaaag
aTCTGGCCCATCAGTTCACACAGGTCCAG[ct]GGGACATGTTCACCCTGGAGGACACG
CTgctaggctaccttgctgatgacctcacatggtgtggtgaattcaacacttccagtga
ggctctgggccctgtgggattgcccagggatgtggagggtgaacagagtgacttctgct
ggaggccctgaatgattagtgtggaggacagagccacaggcacccatcctgatgccatc
tatacttatattagtccatttgtgttgctattaaggaatacctgaggctgcgtaattta
taaagaaaagaggtttatttgactcacagttacgcaggctgtacaagaagtagggtacc
agcatccacttcgggtgaaggcctgaggctgtttccactcatggagaaggggaagggga
gctggcatttacagagatcacatggtgagggaggaaagcaaggagaggtcaggggaggt
gccaggctgtttgtaatgacc

FIGURE 23gg

47878
TCAAATTATCATCGCTTTTTTATTTCAGGATTACACCAAAGACTGTTTCCAACTTGACT
GAGGTAGGTAGTCTTGGATAGACTGGGGGAAATAAGTCCTGTGGGACCTCCTGCCTTAA
AGAAAGCAGGCGGAGGGCCCTAAAGGAAATCAGGCAACCAGACCAAAAGAATGTGGACC
AGGTGGTCCATGCTGTGTCTCTTGTGACCCTTCTTCTCCCTGCCATGTCTTTTGGGAGA
GCCCTTGTGTTGCAAAAATGAGAGTGTGGTGGTATGGATTGGGGTTTAGGCAGAACAGT
ACTGGCCAAGCAGCGCCTCCCTGGACCTCAATTTTCCCTCTGTGGAATGGGCTAGCAAT
CCTGGGCCTCCCCAGGGCGAAGGAAAGACCACTCAGGAAGGGCACCGTCTGGGGCAGGA
AAACGGAGTGGGTTGGATGTATTTTTTTCACGGATGGGCATGAGGATGAATGCTTGTCC
AGGCCGTGCAGCATCTGCCTTGTGGGTCACTTCTGTGCTCCAGGGAGGACTCACCATGG
GCATTTGATTGGCAGAGCAGCTCCGAGTCC[ag]TCCAGAGCTTCCTGCAGTCAATGAT
CACCGCTGTGGGCATCCCTGAGGTCATGTCTCGTAAGTGTGGGCTGGAGGGGAAACTGG
GTGCCGAGGCTGACAGAGCTTCCCATTTCACCTTGTGGGCCCTTCCCAGGCAGAGCTTC
AGGTGCCCCTCTTCCCAGTCATTGATACTTAGCGGTCCTGGCCCCCTTTCCTCTCCCTG
CTGGTGGTATTGCACGCCAATGACTCGGCCAGATGCCCAGACCCCTGTTCTTGGTTTAC
CTGCAGAATATTATCTTTGCCACCCCGCGGGATGGCTCAACCCACTTTCAGGATGCAGG
TCTCCTAATAGCAACCTGATATAGCAGAAAGACCCCTGGGCTGGGAGTCTGAGACCTAG
TTCTAGCCCAGCCCTGAACCTCAGTTTCCCTTTCTGTGAAACAAGAATGTTGAACTTGA
TGATTCCCAATTTTCCTTTTG

tcaaattatcatcgcttttttatttcaggattacaccaaagactgtttccaacttgact
gaggtaggtagtcttggatagactgggggaaataagtcctgtgggacctcctgccttaa
agaaagcaggcggagggccctaaaggaaatcaggcaaccagaccaaaagaatgtggacc
aggtggtccatgctgtgtctcttgtgacccttcttctccctgccatgtcttttgggaga
gcccttgtgttgcaaaaatgagagtgtggtggtatggattggggtttaggcagaacagt
actggccaagcagcgcctccctggacctcaattttccctctgtggaatgggctagcaat
cctgggcctccccagggcgaaggaaagaccactcaggaagggcaccgtctggggcagga
aaacggagtgggttggatgtatttttttcacggatgggcatgaggatgaatgcttgtcc
aggccgtgcagcatctgccttgtgggtcacttctgtgctccagggaggactcaccatgG
GCATTTGATTGGCAGAGCAGCTCCGAGTCC[ag]TCCAGAGCTTCCTGCAGTCAATGAT
CACCGCtgtgggcatccctgaggtcatgtctcgtaagtgtgggctggaggggaaactgg
gtgccgaggctgacagagcttcccatttcaccttgtgggcccttcccaggcagagcttc
aggtgcccctcttcccagtcattgatacttagcggtcctggccccctttcctctccctg
ctggtggtattgcacgccaatgactcggccagatgcccagacccctgttcttggtttac
ctgcagaatattatctttgccaccccgcgggatggctcaacccactttcaggatgcagg
tctcctaatagcaacctgatatagcagaaagacccctgggctgggagtctgagacctag
ttctagcccagccctgaacctcagtttccctttctgtgaaacaagaatgttgaacttga
tgattcccaattttccttttg

FIGURE 23hh

47880
GGGCCAAGGCCACAAAGTCTCAGGACAAGGCAGACTGCAGACCCAGGGGACGTGCGCGG
ACCGGGGCTTGTTTCGGTCCTGGGTGTTCTCAGCCTTGATGTGGACACTAGCGGCTCTG
GTGCACTTGCTCGGAGGAAGCAGCCACGTGTGGGTGTCCTGGCCTCAGCCGGCAGTAAC
CAGCAGACACACAGCACGGAACCCTCCACCCTACCAGGAAGCCCAGGCAAGACCCCCCA
GCAGTGCATGCTGACCCCAGACCCTGGCGACGGATCGGAGCTCCTCGGATTTGGAGTGG
ATCCTTACAAATCCTGCACACTAGACAGCAGACACAGGCCCTGCCAGAGCCAGGGACCC
GAATTTTTGTTTGGAAAAACACTGAGGTAAGTGGGGGGTGGCTCCTGTCCAGGCAGCCC
GGCCGGTGGGACAGTGGGGAGGGTCGGCTCCAAGCCCTCCTGAGCCCTAGAGGGGGTGC
GGGACGGGGACTCACAGGAGATGCAGGACGGCCCGAACATAGTAATTCCTGGTAAAGGG
CCCGAACAGCTTCACCACGGCGGTCATGT [ga] CTTCTGTCCCCTGGGGGAGGGAGGAA
GGCGAGACGGCGCGGCTGGGCCTCTCCCACTCGGGACTCCTTTGCTGCCCTGCTGACCA
CCCCAGGGCACCCAGGCCTCTTTCCTCCCACAAAACACACCGGGCAGGCACCGGCCTTG
GTTTACCCACAAGCACCAAAGGGTTGGTTCCGGAGCCTCCAAGTGAGAAACCAAGCTCC
ACCCAACCCTGTGAGCCCTGCCTGGGCCCCGCAGCCCCCGGAGAGACCCCAGAGCAGGA
GGAGACTCACCAGCGCTCCATGGTGGAGCCCTTCTTCCTCTTCCCCCGGGGGTACTCCA
GCAGGCACACAAACACGCCCGCCACACTGAAGCCATGTGGTTAAGGAACAGCCCAGCTC
AGCCTGAGGGGCCACAGGGAACTCCCTTTACTGAAGACAACACAGAGAGGGGCCCGAGC
ACGGTGGCTCATGCCTGGAAT

gggccaaggccacaaagtctcaggacaaggcagactgcagacccaggggacgtgcgcgg
accggggcttgtttcggtcctgggtgttctcagccttgatgtggacactagcggctctg
gtgcacttgctcggaggaagcagccacgtgtgggtgtcctggcctcagccggcagtaac
cagcagacacacagcacggaaccctccaccctaccaggaagcccaggcaagacccccca
gcagtgcatgctgaccccagaccctggcgacggatcggagctcctcggatttggagtgg
atccttacaaatcctgcacactagacagcagacacaggccctgccagagccagggaccc
gaatttttgtttggaaaaacactgaggtaagtggggggtggctcctgtccaggcagccc
ggccggtgggacagtggggagggtcggctccaagccctcctgagccctagagggggtgc
gggacggggactcacaggagatgcaggacggcccgaacatagtaattcctggtaaaggg
ccCGAACAGCTTCACCACGGCGGTCATGT [ga] CTTCTGTCCCCTGGGGGAGGGAGGAA
Ggcgagacggcgcggctgggcctctcccactcgggactcctttgctgccctgctgacca
ccccagggcacccaggcctctttcctcccacaaaacacaccgggcaggcaccggccttg
gtttacccacaagcaccaaagggttggttccggagcctccaagtgagaaaccaagctcc
acccaaccctgtgagccctgcctgggccccgcagcccccggagagaccccagagcagga
ggagactcaccagcgctccatggtggagcccttcttcctcttcccccgggggtactcca
gcaggcacacaaacacgcccgccacactgaagccatgtggttaaggaacagcccagctc
agcctgaggggccacagggaactccctttactgaagacaacacagagaggggcccgagc
acggtggctcatgcctggaat

FIGURE 23ii

47885
TAATACATGAAAGAAGAAGCTAGTCAATGTGGAGCTCTATTGTGTCCCGGGATCAACAA
AGACAAGATATCTTTAAAATCGTCTTCTAAATTTACCCTAATGTAAAACAAATCCAATA
AAACTCTAATGTAATTTTTTAAGAATTTAAATTTGGAATAATTCCAAAGAACAATTTTT
CTTAATTTTCTACAGCCAGAATATATACCTTTAAAAAAAATGAAAACAGAGATTAACTT
TCTCAGAATTGGTTGACTCACTCTTTCCTTTTATTTTTCTTCCATGGAATTTTCCAGTT
AACTTGAGAAAGTGGAATCGAATTCCGATGTTGAATTTTCCTTCTGGCCCCATTCATGT
GGCAGGTGGTGATTCAGGTACTACTGGGGGCTGCTCAGACAAACCTCCTCATCAGACAT
CAAGAGGCTGTTGCACCAGGAGGGCCGGTACCGTGTCTAGAGGTGGTCGGCATGGGGTT
GGAGTTGTATTACATAAACCCTACTCCAAACAAATGCATGGGGATGTGGCTGGAGTTCC
CCGTTGTCTAACCAGTGCCAAAGGGCAGG[at]CGGTACCTCACCCCACGTTCTTAACT
ATGGGTTGGCAACATGTTCCTGGATGTGTTTGCTGGCACAGTGACAGGTGCTAGCAACC
AGGGTGTTGACACAGTCCAACTCCATCCTCACCAGGTCACTGGCTGGAACCCCTGGGGG
CCACCATTGCGGGAATCAGCCTTTGAAACGATGGCCAACAGCAGCTAATAATAAACCAG
TAATTTGGGATAGACGAGTAGCAAGAGGGCATTGGTTGGTGGGTCACCCTCCTTCTCAG
AACACATTATAAAAACCTTCCGTTTCCACAGGATTGTCTCCCGGGCTGGCAGCAGGGCC
CCAGCGGCACCATGTCTGCCCTCGGAGTCACCGTGGCCCTGCTGGTGTGGGCGGCCTTC
CTCCTGCTGGTGTCCATGTGGAGGCAGGTGCACAGCAGCTGGAATCTGCCCCCAGGCCC
TTTCCCGCTTCCCATCATCGG

taatacatgaaagaagaagctagtcaatgtggagctctattgtgtcccgggatcaacaa
agacaagatatctttaaaatcgtcttctaaatttaccctaatgtaaaacaaatccaata
aaactctaatgtaattttttaagaatttaaatttggaataattccaaagaacaattttt
cttaattttctacagccagaatatatacctttaaaaaaaatgaaaacagagattaactt
tctcagaattggttgactcactctttccttttatttttcttccatggaattttccagtt
aacttgagaaagtggaatcgaattccgatgttgaattttccttctggccccattcatgt
ggcaggtggtgattcaggtactactgggggctgctcagacaaacctcctcatcagacat
caagaggctgttgcaccaggagggccggtaccgtgtctagaggtggtcggcatggggtt
ggagttgtattacataaaccctactccaaacaaatgcatggggatgtggctggAGTTCC
CCGTTGTCTAACCAGTGCCAAAGGGCAGG[at]CGGTACCTCACCCCACGTTCTTAACT
ATGGGTTGGcaacatgttcctggatgtgtttgctggcacagtgacaggtgctagcaacc
agggtgttgacacagtccaactccatcctcaccaggtcactggctggaacccctggggg
ccaccattgcgggaatcagcctttgaaacgatggccaacagcagctaataataaaccag
taatttgggatagacgagtagcaagagggcattggttggtgggtcaccctccttctcag
aacacattataaaaaccttccgtttccacaggattgtctcccgggctggcagcagggcc
ccagcggcaccatgtctgccctcggagtcaccgtggccctgctggtgtgggcggccttc
ctcctgctggtgtccatgtggaggcaggtgcacagcagctggaatctgcccccaggccc
tttcccgcttcccatcatcgg

FIGURE 23jj

47887
AAAAAGAGGAATTAAATTGTGTAGATGCCTTTAAAGAACATTTTTCTAGCATCTTTCTA
CATCTTTCCCTAAGTGGCCTCTTGAGCCCAGTCGGATTTTGGTTATATGCCATGATAGT
AATCATAAGAATCAGTTAAAAATGATCCAAAAATGCACGAATACAGTCGATTCCCTCTC
ATTTATTCCTTGTGGAAAAAGAAAAACACAAATCTTAAAAACTAAAGCAAGTCAGGGAA
GCCTGGAAAGATACCCAGATTTGATAACATGTTAGAAGGAAATCCAGGCTAAGGAATCT
CATTTTCTAGCTTTGATCTGGTTGTCAGTTGGGATGGACTTGCCCAAGTGATGGCCCAC
AGAAAGGCCAAATTTCTTGTTTTTCTCCTCATCCTGTACCTCTTTTTTCATTAAGAATC
CTGCCTGGAAGTTTAGGTCAAAGAGGCTGCTTGGAGCAAAATACAGTGGTGTCTCATCC
CAAATATTCTCCAGGCGTTTCTTCCATCCTTCCAGGATTTGAATTCGGGCGTCTGCTGG
AGTGTGCCCAATGCTATATGTCAGTTGAG[ga]TTCTAAGACTTGGAAGCCACAGAAAT
GCAGAATGCCACTCTGAGGATACAGAAAGCACAGAGAGGTAAGTCAACCAATTCCATGC
AGTTGTACTATAAACAACAGAAGTTGGTCTGGGCTTCTCAGTAAGACACTCTGATAAGG
AGGCCTCAGGCACACTAGAGAATCAGTTCAGAGCTAGCGTCTCTCTCTTACCCTCTACC
TAGCCGTTACCAATTTTAGCCTTCTCAGGTGTGTTCTTCTTTAAATGCATAAACCTTGA
AACTGTGCCAACCTGGATCCTTTGCCAAGAAGGCTGGAAGTTCTGTTACTTTAGGGAGT
CTCAGTTTCTTGGCAGGTGACTCACCAAGACCTGCGTGGGTGCATTTCTCTGCCTCTCC
ATATAACTAGATGAGTCCTTTTTTTCTTTTTCTTTTTTTTTTTTTTTTGAGGCAGAGT
CTCGCTCGGTCGTCCAGGCTG

aaaaagaggaattaaattgtgtagatgcctttaaagaacatttttctagcatctttcta
catctttccctaagtggcctcttgagcccagtcggattttggttatatgccatgatagt
aatcataagaatcagttaaaaatgatccaaaaatgcacgaatacagtcgattccctctc
atttattccttgtggaaaaagaaaaacacaaatcttaaaaactaaagcaagtcagggaa
gcctggaaagatacccagatttgataacatgttagaaggaaatccaggctaaggaatct
cattttctagctttgatctggttgtcagttgggatggacttgcccaagtgatggcccac
agaaaggccaaatttcttgtttttctcctcatcctgtacctcttttttcattaagaatc
ctgcctggaagtttaggtcaaagaggctgcttggagcaaaatacagtggtgtctcatcc
caaatattctccaggcgtttcttccatccttccaggatttgaattcgggCGTCTGCTGG
AGTGTGCCCAATGCTATATGTCAGTTGAG[ga]TTCTAAGACTTGGAAGCCACAGAAAT
GCAGAATGCCACTctgaggatacagaaagcacagagaggtaagtcaaccaattccatgc
agttgtactataaacaacagaagttggtctgggcttctcagtaagacactctgataagg
aggcctcaggcacactagagaatcagttcagagctagcgtctctctcttaccctctacc
tagccgttaccaattttagccttctcaggtgtgttcttctttaaatgcataaaccttga
aactgtgccaacctggatcctttgccaagaaggctggaagttctgttactttagggagt
ctcagtttcttggcaggtgactcaccaagacctgcgtgggtgcatttctctgcctctcc
atataactagatgagtcctttttttctttttcttttttttttttttttgaggcagagt
ctcgctcggtcgtccaggctg

FIGURE 23kk

47889
CTCATGTAGGAACCAGCAGGCTACTAGAAATTAAAGTTTAAATCTGGGAGAAGGTTAGC
GTTAAGTGTGTGTATACGAGAGTCAAGCCAGAGAGGAGGGCAGTAATGCTGTGGGGTTG
CATGAAATTCACCAAAGGAGAGCATGCAAAGTGAAGAGGGAGGCAAATGAAGATGGAGC
CCCAAGGAGCACTTACATTTAAAAATATGGGCAGAGGAAGAGGAATCGTCAAAGGAGAC
TAAAAAGTAGCCAAGGCAGGGAGCATTTCAAGAAGGAGAGAAAGATCCACTTTGCCATA
TGCTGCAGAAAGAGTCCAACAGGTTGAGAAATGACAGTACTCGTGATTCCAAAGGTAAT
GAAAAAAATCCCCAGAATTCTATGCATGAATTAATTACGTGATTAAACATACAAATGTA
CTGTTCTCCAAGAAAACTGAGCTGTTTCCATATTCAGCATTGAATACCAAGATATTATT
TTCTTGTTTGTAGAGATATTCATGATCTAAAGAGAGAAAACACCCAGATCAAAATTTCA
AGTTGTTATTAAACATCTTCATAAGCTGA[ga]AATTACAGAATACAGTTTAAGCTCAC
AAATACCAAATAGGCATTTCTAAGTTGAGAAAACATGAATGATATTATACTAACATTCA
TTCATTTTTTCATCATTATTGTCAAGGTTTCAATTCACATTTAATTTTTTATTATACAT
GTCAAAGAAATACTTGGGTTCCTTTCAGTCTTTCTCCCTTTGCACTTCAAGTAGAAAAA
GAAAAAAAAAACTCTCTATAGAATTTTTAAAAACAAGGATTACCTCTTCTCAGTGCCAT
AAAAGCCCACATCTCGACTTAACTAGAATGAATGTAAGCATAAAATCTGCCCTACCCCA
AAAAATTCTTACCTGAAATCCATCTTAAGGAGTATAACTTCAGTCTATAAGTATTTTTT
AAGTAATCAGTTAGAGTGTAAGTTTTGCGACTGTCAGCTGTAGCATCATCTGCTGGTTG
AAAGAAAGAGCCAAATGTTCA

ctcatgtaggaaccagcaggctactagaaattaaagtttaaatctgggagaaggttagc
gttaagtgtgtgtatacgagagtcaagccagagaggagggcagtaatgctgtggggttg
catgaaattcaccaaaggagagcatgcaaagtgaagagggaggcaaatgaagatggagc
cccaaggagcacttacatttaaaaatatgggcagaggaagaggaatcgtcaaaggagac
taaaaagtagccaaggcagggagcatttcaagaaggagagaaagatccactttgccata
tgctgcagaaagagtccaacaggttgagaaatgacagtactcgtgattccaaaggtaat
gaaaaaaatccccagaattctatgcatgaattaattacgtgattaaacatacaaatgta
ctgttctccaagaaaactgagctgtttccatattcagcattgaataccaagatattatt
ttcttgtttgtagagatattcatgatctaaagagaGAAAACACCCAGATCAAAATTTCA
AGTTGTTATTAAACATCTTCATAAGCTGA[ga]AATTACAGAATACAGTTTAAGCTCAC
AAATACCAAATAGGCATTTCTAAGTTGagaaaacatgaatgatattatactaacattca
ttcattttttcatcattattgtcaaggtttcaattcacatttaattttttattatacat
gtcaaagaaatacttgggttcctttcagtctttctccctttgcacttcaagtagaaaaa
gaaaaaaaaaactctctatagaatttttaaaaacaaggattacctcttctcagtgccat
aaaagcccacatctcgacttaactagaatgaatgtaagcataaaatctgccctacccca
aaaaattcttacctgaaatccatcttaaggagtataacttcagtctataagtatttttt
aagtaatcagttagagtgtaagttttgcgactgtcagctgtagcatcatctgctggttg
aaagaaagagccaaatgttca

FIGURE 2311

47892
AGCGTTCAGAGAAGGAGCGCAGGCAGAAGTCACCGCGGGCGGCGGAGACGCGCGTCCTG
CACCGCTGCTCCGGGCGGTGGAGTCACTCGCCGCTGGCAAGTTTCGGCCCCGAGTTAAA
CATTAGTGAGCGCCGAGCCCGCTGGGTATAAAGGCGCCGCGGGCAGGCTGCAGGGCAGG
CGGCGCGGGAGCAGGCGCGCGTGGCGCGGGGCACTGGCATCCCGGCCGGGGGGAGCCCG
CGAGGGCCCCCTGAGGGCGGTGTAGGGCGCTGGGCGGCAGCCGGGGCGCAGAGTGCGGG
GCCCGGAGGAGCCGTGGGGGAGGGGAAAGGCGCGCGGCCTCGGATGCGCAGACCCTGG
GCCGGCGACTCGGGGACCCTGCTCCCTCTTAGCTAAAAATGACGTCGGCGTTCAGCTCC
TCCAACCTCACGTGGACAGGCGAGGGAACCGAGACCCAGAGAGGGCAGGGGACTTTGGC
AAACTCACACAGCCCACCGCAGGCAACTGGAACTGAAACCCAGGACTCCGTCTCTTGCC
AGTGAAAGTTATGTTAGGAAGCAGTGAGG[tg]GTCTAAAGCAGTATGAAAGGCAAAGA
GAAAAGGTGATTGTTCCCTCTTGAATGGCCCTTGGAAGCTGAGTATCTGGATTCACCCT
CCCTAGGGAATTTCCCGATTGTCTTGCAGGCTTACACACTCATCAAGATGACAAAAATA
ATGACAGTAACACTTATGTGGAACTTGACTTTTTCCCAGGTGCTGCTCTAAGCATTTAC
TGTGTTTGTTTTACAGGAAGGAAGACTGTACACAGAGAATAAATAACTTGGCCAAGCCA
TTCAGCTAGGAAGTTGTAGATCCTAAATTAAGAGTTCAAGGTCTTAATGGCTACTCTAT
GCGGCCTCTCATAGTCTTTTCAAGGGTTTTGGAGAAGAATAAAAGATCAGGTATGGCTT
CTCCCTCCCCCAGCTCTCTATTGTTCCCTAAAGGATTATTCATTCGTTCATTCATTCCT
ACATCCTCCCATTTATTCCAG

agcgttcagagaaggagcgcaggcagaagtcaccgcgggcggcggagacgcgcgtcctg
caccgctgctccgggcggtggagtcactcgccgctggcaagtttcggccccgagttaaa
cattagtgagcgccgagcccgctgggtataaaggcgccgcgggcaggctgcagggcagg
cggcgcgggagcaggcgcgcgtggcgcggggcactggcatcccggccggggggagcccg
cgagggccccctgagggcggtgtagggcgctgggcggcagccggggcgcagagtgcggg
gcccggaggagccgtgggggaggggaaagggcgcgcggcctcggatgcgcagaccctgg
gccggcgactcggggaccctgctccctcttagctaaaaatgacgtcggcgttcagctcc
tccaacctcacgtggacaggcgagggaaccgagacccagagagggcaggggactttggc
aaactcacacagcccaccgcaggcaactggaactgaaacccAGGACTCCGTCTCTTGCC
AGTGAAAGTTATGTTAGGAAGCAGTGAGG[tg]GTCTAAAGCAGTATGAAAGGCAAAGA
GAAAAGGTGATTGTTCCCTCTtgaatggcccttggaagctgagtatctggattcacccct
ccctagggaatttcccgattgtcttgcaggcttacacactcatcaagatgacaaaaata
atgacagtaacacttatgtggaacttgacttttttcccaggtgctgctctaagcatttac
tgtgtttgttttacaggaaggaagactgtacacagagaataaataacttggccaagcca
ttcagctaggaagttgtagatcctaaattaagagttcaaggtcttaatggctactctat
gcggcctctcatagtcttttcaagggttttggagaagaataaaagatcaggtatggctt
ctccctcccccagctctctattgttccctaaaggattattcattcgttcattcattcct
acatcctcccatttattccag

FIGURE 23mm

47895
TTCTGATCAGTTTTCTATGTTAAATAAATATACATCTACCTTGTCAGTTTAGATGACTG
TACTGGACTCCAGTATACTGTCAAACTATACTTGATTAATCCTGTATTGCTGGATACGT
GGGGCTTTCTCCCTACCCTCCAGATTTTAAATTATTGAACAAGTATTTATGGAGGCCTG
CTGTGAGCCAGGAGCTGTCCTGAGCCCTGGAAACCCAGCAGTGGCTGTACAGACCTGGC
CCAGCTGTCAGGGGGCACCTCTAAGGAAACCGGGAGGCAATAATCGTAGCTCCCTTGCA
GGGAGGTTGTGAAGGCTGAGTGAGGACATCTGTGCACCTGGAGCACAGTGTGAGTGTGA
AACCAGTGTCAGCCCTTATTACTGTCAATACCATGAAGGGGCGGCGGGGGCACTAAGGG
TGGCAGGACTCAATATCTAGGCTCTGGGGGGTGCCAGAGCCTGACCGTGCAGGGTCTTC
TCTCTCCCTCCACCCTGACTGTGCTCTGTCCCCCCAGGGCTGGACATCCACTTCATCCA
CGTGAAGCCCCCCCAGCTGCCCGCAGGCC[ag]TACCCCGAAGCCCTTGCTGATGGTGC
ACGGCTGGCCCGGCTCTTTCTACGAGTTTTATAAGATCATCCCACTCCTGACTGACCCC
AAGAACCATGGCCTGAGCGATGAGCACGTTTTTGAAGTCATCTGCCCTTCCATCCCTGG
CTATGGCTTCTCAGAGGCATCCTCCAAGAAGGGTACGGGGCTGCTAGAGGTTCCATAAC
TGCCCCGTCCTCGCCAAGGGTGGGCCCGGTGTTCCCACCAGGCTCTCCTTCCGGCGGGG
TGAGCAGGGAGTTGGCCCGAGGAAGCTGGGAAAGGAGGGGCCTGAGAGGCCGGCCCCAG
ACACACCGCCCTCCGGGGCTGGAGATGCCACCCCTATATTTGGGCTCCAGGATTCCTTC
TTGCCTCTGTGAGCTTTTCTGACCTCCACCTGGGGGTAGGCGGGCCTGAGAAATTTCAT
AGAACACCAGAGGGCCCAAGG

ttctgatcagttttctatgttaaataaatatacatctaccttgtcagtttagatgactg
tactggactccagtatactgtcaaactatacttgattaatcctgtattgctggatacgt
ggggctttctccctaccctccagattttaaattattgaacaagtatttatggaggcctg
ctgtgagccaggagctgtcctgagccctggaaacccagcagtggctgtacagacctggc
ccagctgtcagggggcacctctaaggaaaccgggaggcaataatcgtagctcccttgca
gggaggttgtgaaggctgagtgaggacatctgtgcacctggagcacagtgtgagtgtga
aaccagtgtcagcccttattactgtcaataccatgaaggggcggcgggggcactaaggg
tggcaggactcaatatctaggctctggggggtgccagagcctgaccgtgcagggtcttc
tctctccctccaccctgactgtgctctgtccccccagggctggacatccacttcatcca
cgtGAAGCCCCCCCAGCTGCCCGCAGGCC[ag]TACCCCGAAGCCCTTGCTGATGGTGC
acggctggcccggctctttctacgagttttataagatcatcccactcctgactgacccc
aagaaccatggcctgagcgatgagcacgtttttgaagtcatctgcccttccatccctgg
ctatggcttctcagaggcatcctccaagaagggtacggggctgctagaggttccataac
tgccccgtcctcgccaagggtgggcccggtgttcccaccaggctctccttccggcgggg
tgagcagggagttggcccgaggaagctgggaaaggaggggcctgagaggccggccccag
acacaccgccctccggggctggagatgccacccctatatttgggctccaggattccttc
ttgcctctgtgagcttttctgacctccacctgggggtaggcgggcctgagaaatttcat
agaacaccagagggcccaagg

FIGURE 23nn

47897
TTTCTGGATCACGTTTTCATATATTCTGGTTCAGTACATCTATCTTTGAGTTATCTTTA
ATATACTGAACCAGAACATACAGGAATGTGATCCAGAACATCATTGGCCATCAGATTTT
CTAGTATATGTGATGTGCACCTCTTATAAATTATAATTGAATTCACTGCCATATCTCCA
AGGGGTGTCACTCTTGTACTCCAGAAGATACTGGTTATGCACAAGAAATCATGCAGGGA
CAAATAGACAGATACCATTTAGTGTTTTGATTTATTCTGAGGGAATTTTAAATTTGTAA
TATGTATCTTAATCATTAAATATTTTTCTTAACCCACTTTTCTTTTTTCATACTGTATC
TGCCAAAACCATTTGCTAGCATAGAAAAGAGGGATTTCTTTCTGTATTTCTCTTAGACA
TTTGTATCCAGTGTAAATAAACATCCTGATTTTGCAACTACTGGCCAGTGGGATGTTAC
CACTGAAAGGGATGGTAAAAAAGAATCGGCTGTCTTTGATGCTGTAATGGTTTGTTCCG
GACATCATGTGTATCCCAACCTACCAAAA[ag]AGTCCTTTCCAGGTAAGGCCAAAATT
TAAGCTGCTAGCCACATAACTGACAAAAATGAATATCTTGATAATGTCTTCTTTTTTCT
AAAAGTATAAGCAGGTTAAATTAAAATATACTTCTGTTATATCTAATATGCTTGGTGTG
TTAAAATAGCACATTATTGTGACTGCATCTATTCACAAGGTCGCTTCTGTTAAAGTCTT
TGTTTAAATATATGACTCAAACTGCCATGTATTTCTCACTTTTCACTCAGGACTAAACC
ACTTTAAAGGCAAATGCTTCCACAGCAGGGACTATAAAGAACCAGGTGTATTCAATGGA
AAGCGTGTCCTGGTGGTTGGCCTGGGGAATTCGGGCTGTGATATTGCCACAGAACTCAG
CCGCACAGCAGAACAGGTACTACTCCCCGGGTACTCGGGTGACTCTCGTTACTGACAGA
AGAGTTATTATCGTTTGAAAG

tttctggatcacgttttcatatattctggttcagtacatctatctttgagttatcttta
atatactgaaccagaacatacaggaatgtgatccagaacatcattggccatcagatttt
ctagtatatgtgatgtgcacctcttataaattataattgaattcactgccatatctcca
aggggtgtcactcttgtactccagaagatactggttatgcacaagaaatcatgcaggga
caaatagacagataccatttagtgttttgatttattctgagggaattttaaatttgtaa
tatgtatcttaatcattaaatatttttcttaacccacttttcttttttcatactgtatc
tgccaaaaccatttgctagcatagaaaagagggatttctttctgtatttctcttagaca
tttgtatccagtgtaaataaacatcctgattttgcaactactggccagtgggatgttac
cactgaaagggatggtaaaaaagaatcggctgtctttgatgctgtaatgGTTTGTTCCG
GACATCATGTGTATCCCAACCTACCAAAA[ag]AGTCCTTTCCAGGTAAGGCCAAAATT
TAAGCTGCTAGCCacataactgacaaaaatgaatatcttgataatgtcttcttttttct
aaaagtataagcaggttaaattaaaatatacttctgttatatctaatatgcttggtgtg
ttaaaatagcacattattgtgactgcatctattcacaaggtcgcttctgttaaagtctt
tgtttaaatatatgactcaaactgccatgtatttctcacttttcactcaggactaaacc
actttaaaggcaaatgcttccacagcagggactataaagaaccaggtgtattcaatgga
aagcgtgtcctggtggttggcctggggaattcgggctgtgatattgccacagaactcag
ccgcacagcagaacaggtactactccccgggtactcgggtgactctcgttactgacaga
agagttattatcgtttgaaag

FIGURE 23oo

47899
TCAAAGAAAATCCAACATTAAAATGTATGCCTTACGATAGGCTTGTGTTCTTATTTGCT
GCCTTCTCTCTCTATGCTGTGCAGCTAGGCTGTAATTTTAAATGCATGTCTTGGATTTT
ATTCTACAAGAAAAGGAATGCATCTGTTTCCATTCCTTACCCTTGGCTGGGGGATAATT
TTAATGTTGGGTTTGAACCCCACGAAAGAATGTTATATTTGCTCTATCTTTTGGTAGAA
ATTAGATTGGTAACCTCGTAGGTCCACAAAAGTAAACTTTCACTTTAAGGGAAAATGAG
TAAGCAAGTAAATATTGCTAGGACTACCACTGGAAAATAATTTAAAGGCTATGTCACA
CTGGAGGTTGGGTAAGTGGTTTAGAGGGGTGCGGGTTAAGACATTCGGGGGCATAATAC
TAAGGAGAGCATCCCCAACCCTAAACATCTTCAAAATGATCAGGGCTTATGGGCACTAT
TTGACGAGCATAAGAACTTAATAATGTCAAGAGAAATTTTAGACCTATTTAATACATTT
ATAAGCAAGTTTTGAGCCAGGCTTAGACT[ct]TTACCTGTTCCTCTTGGTATTCATCA
ACCACTGCACAAAATCTTGGGCACGCCTGGAGTCCAGATACTTGCTGTAGTCACTGGTG
AATGTGCCCTGTGAATGGCGCTTGTCCTCGTTCATCTGATCAGGATCACTGAGTGGGTC
TGCCTGGGAAGCTGAGAATGATCTGTGAAGAACAGTGATTGGTACAACATAAATCTCTC
CTCAAGAGTAGACTCACTTGAGAAGCATCTTCACTACAAAATACAAGACCATATAAAAC
AGTAAGGCAGGCATCTAGAGTATTTCAATAGGTAGTTTAGAAAGATCTTCCTTAGCTTG
TCATGAGAATCCCTTCGTTTTAGTATAGTTGCATACGCTATTATTCTGAATTCTAGAAA
CATGTTTCTCAACTGACTTCTTTTTTTCTGAAATAGGATTAAACAAATCTTTTTCTACT
AATTAATCTACTCATGATTAT

tcaaagaaaatccaacattaaaatgtatgccttacgataggcttgtgttcttatttgct
gccttctctctctatgctgtgcagctaggctgtaattttaaatgcatgtcttggatttt
attctacaagaaaaggaatgcatctgtttccattccttacccttggctgggggataatt
ttaatgttgggtttgaaccccacgaaagaatgttatatttgctctatcttttggtagaa
attagattggtaacctcgtaggtccacaaaagtaaactttcactttaagggaaaatgag
taagcaagtaaatattgctaggactaccactggaaaataatttaaaggctatgtcaca
ctggaggttgggtaagtggtttagaggggtgcgggttaagacattcgggggcataatac
taaggagagcatccccaaccctaaacatcttcaaaatgatcagggcttatgggcactat
ttgacgagcataagaacttaataatgtcaagagaaattttagaCCTATTTAATACATTT
ATAAGCAAGTTTTGAGCCAGGCTTAGACT[ct]TTACCTGTTCCTCTTGGTATTCATCA
ACCACTGCACAAAATCTTGggcacgcctggagtccagatacttgctgtagtcactggtg
aatgtgccctgtgaatggcgcttgtcctcgttcatctgatcaggatcactgagtgggtc
tgcctgggaagctgagaatgatctgtgaagaacagtgattggtacaacataaatctctc
ctcaagagtagactcacttgagaagcatcttcactacaaaatacaagaccatataaaac
agtaaggcaggcatctagagtatttcaataggtagtttagaaagatcttccttagcttg
tcatgagaatcccttcgttttagtatagttgcatacgctattattctgaattctagaaa
catgtttctcaactgacttctttttttctgaaataggattaaacaaatctttttctact
aattaatctactcatgattat

FIGURE 23pp

47901

TCTGCCTGTCCGTCTGCCTGTCTGTCTGCCTGTCCATCTGTCCATCTGCCTATCCATCT
GCCTGCCTGTCTGTCGGCCTGCCTGCCTGCCTGTCTGTCTGCTGCCTGTCTGTCCGTCT
GCCTGTCTGCCTGTCCGTCTGCCTGCCTGTCCGTCTGCCTGTCCGTCTGCCTGCCTGCC
TGTCTGTCTGCCTGCCTGTCTGCCTGCCTGTCCGTCTGCCTGTCCGTCTGCCTGCCTGT
CTGCCTGCCTGTCTGCCTGTCTGCCCGTCTGCCTGTCTGTCTGCCTGTCCGTCTGCCTG
TCTGTCCGTCTGTCCATCTGCCTATCCATCTGCCTGCCTATCTGTCTGTCCGTCTGCCT
GCCTGTCTGTCTGCCTGTCTGCCTGTCTGTCTGCCTGTCTGTCCATCTGCCTATCCATC
TACCTGCCTGCCTGTCTGCCTGTCTGTCTGCCTGTCTGTCTGCCTGCCTGTCTGTCTGT
CTGTCTGGTTGCTTGTGCATGTGTCCCCCAGCCACAGGTCCCCTCCGCTCAGGTGATGG
ACTTCCTGTTTGAGAAGTGGAAGCTCTAC [ag] GTGACCAGTGTCACCACAACCTGAGC
CTGCTGCCCCCTCCCACGGGTGAGCCCCCCACCCAGAGCCTTTCAGCCTGTGCCTGGCC
TCAGCACTTCCTGAGTTCTCTTCATGGGAAGGTTCCTGGGTGCTTATGCAGCCTTTGAG
GACCCCGCCAAGGGGCCCTGTCATTCCTCAGGCCCCCACCACCGTGGGCAGGTGAGGTA
ACGAGGTAACTGAGCCACAGAGCTGGGGACTTGCCTCAGGCCGCAGAGCCAGGAAATAA
CAGAACGGTGGCATTGCCCCAGAACCGGCTGCTGCTGCTGCCCCCAGGCCCAGATGGGT
AATACCACCTACAGCCCCGTGGAGTTTTCAGTGGGCAGACAGTGCCAGGGCGTGGAAGC
TGGGACCCAGGGGCCTGGGAGGGCTCGGGTGGAGAGTGTATATCATGGCCTGGACACTT
GGGGTGCAGGGAGAGGATAGG

tctgcctgtccgtctgcctgtctgtctgcctgtccatctgtccatctgcctatccatct
gcctgcctgtctgtcggcctgcctgcctgcctgtctgtctgctgcctgtctgtccgtct
gcctgtctgcctgtccgtctgcctgcctgtccgtctgcctgtccgtctgcctgcctgcc
tgtctgtctgcctgcctgtctgcctgcctgtccgtctgcctgtccgtctgcctgcctgt
ctgcctgcctgtctgcctgtctgcccgtctgcctgtctgtctgcctgtccgtctgcctg
tctgtccgtctgtccatctgcctatccatctgcctgcctatctgtctgtccgtctgcct
gcctgtctgtctgcctgtctgcctgtctgtctgcctgtctgtccatctgcctatccatc
tacctgcctgcctgtctgcctgtctgtctgcctgtctgtctgcctgcctgtctgtctgt
ctgtctggttgcttgtgcatgtgtcccccagccacaggtcccctccgctcaggtgatgG
ACTTCCTGTTTGAGAAGTGGAAGCTCTAC [ag] GTGACCAGTGTCACCACAACCTGAGC
CTGCtgccccctcccacgggtgagccccccacccagagcctttcagcctgtgcctggcc
tcagcacttcctgagttctcttcatgggaaggttcctgggtgcttatgcagcctttgag
gaccccgccaaggggccctgtcattcctcaggcccccaccaccgtgggcaggtgaggta
acgaggtaactgagccacagagctggggacttgcctcaggccgcagagccaggaaataa
cagaacggtggcattgccccagaaccggctgctgctgctgcccccaggcccagatgggt
aataccacctacagccccgtggagttttcagtgggcagacagtgccagggcgtggaagc
tgggacccaggggcctgggagggctcgggtggagagtgtatatcatggcctggacactt
ggggtgcagggagaggatagg

FIGURE 23qq

47905
CTGCTTTCCAAATCAGCTTGGAGAGACAGGCTGACTCCTTTCCCTCTTCCTCAGGCATC
CTCTCTGGCCACGATAACAGGGTGAGCTGCCTGGGAGTCACAGCTGACGGGATGGCTGT
GGCCACAGGTTCCTGGGACAGCTTCCTCAAAATCTGGAACTGAGGAGGCTGGAGAAAGG
GAAGTGGAAGGCAGTGAACACACTCAGCAGCCCCCTGCCCGACCCCATCTCATTCAGGT
GTTCTCTTCTATATTCCGGGTGCCATTCCCACTAAGCTTTCTCCTTTGAGGGCAGTGGG
GAGCATGGGACTGTGCCTTTGGGAGGCAGCATCAGGGACACAGGGGCAAAGAACTGCCC
CATCTCCTCCCATGGCCTTCCCTCCCCACAGTCCTCACAGCCTCTCCCTTAATGAGCAA
GGACAACCTGCCCCTCCCCAGCCCTTTGCAGGCCCAGCAGACTTGAGTCTGAGGCCCCA
GGCCCTAGGATTCCTCCCCCAGAGCCACTACCTTTGTCCAGGCCTGGGTGGTATAGGGC
GTTTGGCCCTGTGACTATGGCTCTGGCAC[ct]ACTAGGGTCCTGGCCCTCTTCTTATT
CATGCTTTCTCCTTTTTCTACCTTTTTTCTCTCCTAAGACACCTGCAATAAAGTGTAG
CACCCTGGTACATCTGTGATGTTTGCCTTCTACTCTCTTCTGTTCCAAAAAGACCCAGG
TCCCATTTAAGGGCAGTAATGTGTTACAGGTGCTGTGATAAAGGCTGGTACTGGATAG
CTTGTGGGCTTATGGGAGGAGGCCTGAGATGGGTCAGGGGGAGAAGGTATTCAGCAGGT
GGCTGGGGGACTGTGTGCAGCAGTTCGCTATGGCCTGCCTGTGGTGCCCATGTGTTTGT
ACGGGAGGGTTAGCTTGAGAAGGAATCAGATTATAAAAGGTCTTGAATGTCAAGCCAGA
GAGTCCAGACTTTTTCCTAAGGGCAATGAGAAGCCATTGAGGAGTTCTGAGCAGAGTAG
TAACATGATCAGTTATGCTTC

ctgctttccaaatcagcttggagagacaggctgactcctttccctcttcctcaggcatc
ctctctggccacgataacagggtgagctgcctgggagtcacagctgacgggatggctgt
ggccacaggttcctgggacagcttcctcaaaatctggaactgaggaggctggagaaagg
gaagtggaaggcagtgaacacactcagcagccccctgcccgaccccatctcattcaggt
gttctcttctatattccgggtgccattcccactaagctttctcctttgagggcagtggg
gagcatgggactgtgcctttgggaggcagcatcagggacacaggggcaaagaactgccc
catctcctcccatggccttccctccccacagtcctcacagcctctcccttaatgagcaa
ggacaacctgcccctccccagccctttgcaggcccagcagacttgagtctgaggcccca
ggccctaggattcctcccccagagccactacctttgtccaggcctgggtggTATAGGGC
GTTTGGCCCTGTGACTATGGCTCTGGCAC[ct]ACTAGGGTCCTGGCCCTCTTCTTATT
CATGCTTTCTCcttttttctaccttttttctctcctaagacacctgcaataaagtgtag
caccctggtacatctgtgatgtttgccttctactctcttctgttccaaaaagacccagg
tcccatttaagggcagtaatgtgttacaggtgctgtgataaaggctgggtactggatag
cttgtgggcttatgggaggaggcctgagatgggtcagggggagaaggtattcagcaggt
ggctgggggactgtgtgcagcagttcgctatggcctgcctgtggtgcccatgtgtttgt
acgggagggttagcttgagaaggaatcagattataaaaggtcttgaatgtcaagccaga
gagtccagactttttcctaagggcaatgagaagccattgaggagttctgagcagagtag
taacatgatcagttatgcttc

FIGURE 23rr

47906
ACCATTGTTCCTGTTTTGCAAAGAAGGCAACAGGCTCAGAGAAGGCCAGTGCCTCGCCC
CAAGACATGCTAGCTCTGACTAGGATGCCATGACCACGCTGTCCCCTGCCCACTACACT
CACCCGGTGTGTAGCCCCAAGGCTCATAGTAGGAGGGGAAGACTCCAAGGTGACAGCCA
CGGACAAACTCCTCATAGTCCACAGGGAGCAGGGGGCTTGTGGAGGAGAGGAACTCCGG
GTGGAAAATCACCTGGTAGTGAAAAAGAAGGACTCAGCCCAAGTGCCTTATTTAGCTAA
GCCCTGAGATCCCAAGGTGGCCCAGAGAGGGTAAAAAGCTTGTCTAGCATCACACAGCA
TGTGTTTGGCAGGACCAATGTTCAAACCCAGGTCTGCCTGCCTCAGAAGCCAGGGTTCT
TTCTAACCACAGCAATACCTTTGATAAAACTTATAGGGGAATGGAGTGTGTGAGGCCCA
GGACCCAACCCCTTCCCTCTGCCGTGCCCAACCCAGCCCTGACCAAATGCCCTCACCTT
CACCCTGTCGGCACTGCTATTGAAGAGGC[tc]GATTCGGCGGATGGTGGTCAGGATGG
GGTCTGAGGAGTCATCCAGCATATTGTGGGTGCACACAGGGGGGAAAGACTGCCGCTGC
AGGAGCCACAAGAAGGGTAAGGGGTCATGGAAGGGACAGAGAACTCCCTACTTCCTCAT
GAGCCATGCGGACCCTGGGGGAGCCAAGGAGACCACAAATGCACCGGACGTGGGGCAAC
AAACCCAAGTGATCACCAGGAGTTGTGGATTCCCACTAGTACAACCTGTAAAGGTTTTC
TTTCTTTTCTTTTAAATTATTATTATTTATTTTTGAGGCGGAGTCTCGCTCTGTCGCCC
AGGCTGGAATGCAGTGGCACAATCTCGGCTCACTGCAAGCTCCACCTCCCAGGATCATG
CCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGAACTACAGGCGCCTACCACCACGCCC
GGTTAATTTTTTGTATTTTTA

accattgttcctgttttgcaaagaaggcaacaggctcagagaaggccagtgcctcgccc
caagacatgctagctctgactaggatgccatgaccacgctgtcccctgcccactacact
cacccggtgtgtagccccaaggctcatagtaggaggggaagactccaaggtgacagcca
cggacaaactcctcatagtccacagggagcagggggcttgtggaggagaggaactccgg
gtggaaaatcacctggtagtgaaaaagaaggactcagcccaagtgccttatttagctaa
gccctgagatcccaaggtggcccagagagggtaaaaagcttgtctagcatcacacagca
tgtgtttggcaggaccaatgttcaaacccaggtctgcctgcctcagaagccagggttct
ttctaaccacagcaatacctttgataaaacttataggggaatggagtgtgtgaggccca
ggacccaaccccttccctctgccgtgcccaacccagccctgaccaaatgccctcacctt
caCCCTGTCGGCACTGCTATTGAAGAGGC[tc]GATTCGGCGGATGGTGGTCAGGATGG
Ggtctgaggagtcatccagcatattgtgggtgcacacaggggggaaagactgccgctgc
aggagccacaagaagggtaaggggtcatggaagggacagagaactccctacttcctcat
gagccatgcggaccctgggggagccaaggagaccacaaatgcaccggacgtggggcaac
aaacccaagtgatcaccaggagttgtggattcccactagtacaacctgtaaaggttttc
tttcttttcttttaaattattattatttatttttgaggcggagtctcgctctgtcgccc
aggctggaatgcagtggcacaatctcggctcactgcaagctccacctcccaggatcatg
ccattctcctgcctcagcctcccgagtagctggaactacaggcgcctaccaccacgccc
ggttaatttttgtattttta

FIGURE 23ss

47909
CAAACTCACAGTTGGATGGCACAACAATTACATCCTGTGTGGTCAGCAGTGATGGAGGG
GCCGCAGAGATTTGGAAACAGGAGGGACACAGGATACAGATATAGGAGGGTAGAAGGCA
GACTTCCTGGAGGAGGTGAAACCTAACCTGAGTCCCTAAGGTGATAGGAAGCAGGAACC
AGGGAAGGGGAGCCTATTCTAACACAGTAGAAGCAGCAACTGCTGAGGTCTGGATGAGG
GGACCTCAACTGTGGCCCAAAACCCCAAGTTCCCATTGTGGCTCTGCCAACAACTGGCT
GTGCGACCCAGGACAAGTCCTATCTTTGCACTGTGTCTGGGTTTCCCCGTGTGTAAGAT
GAGGCGGTTGCTAGGTGCTTATTGGATGCATTCCTCAAGTCCCGCCCTCCATCTCCTAT
TCCCCTCTCTTCTGGTTTAGTGCTTTAGGAAATGTGGCAGAAATCTTTTTCTGCCTGTG
TCTAGGAAATCATAATTCATGCTGGCGTACCCTGGTTGTTGAGGTCCCTGAATCCTTGT
GCCCACACTGCTGAAGACTCCTTGTGTGA[ct]ACAAGTCAGGGGACATCTGGGTCTTG
ACTCCCCAGATGCTCCAGCTGGACCCTGCTGCCCTCCCTTGCCCACCCTCTTCCATTGT
AGATGCCAAGGGGCTGAGCGATCCAGGGAAGATCAAGCGGCTGCGTTCCCAGGTGCAGG
TGAGCTTGGAGGACTACATCAACGACCGCCAGTATGACTCGCGTGGCCGCTTTGGAGAG
CTGCTGCTGCTGCTGCCCACCTTGCAGAGCATCACCTGGCAGATGATCGAGCAGATCCA
GTTCATCAAGCTCTTCGGCATGGCCAAGATTGACAACCTGTTGCAGGAGATGCTGCTGG
GAGGTCCGTGCCAAGCCCAGGAGGGGCGGGGTTGGAGTGGGGACTCCCCAGGAGACAGG
CCTCACACAGTGAGCTCACCCCTCAGCTCCTTGGCTTCCCCACTGTGCCGCTTTGGGCA
AGTTGCTTAACCTGTCTGTGC

caaactcacagttggatggcacaacaattacatcctgtgtggtcagcagtgatggaggg
gccgcagagatttggaaacaggagggacacaggatacagatataggagggtagaaggca
gacttcctggaggaggtgaaacctaacctgagtccctaaggtgataggaagcaggaacc
agggaaggggagcctattctaacacagtagaagcagcaactgctgaggtctggatgagg
ggacctcaactgtggcccaaaaccccaagttcccattgtggctctgccaacaactggct
gtgcgacccaggacaagtcctatctttgcactgtgtctgggtttccccgtgtgtaagat
gaggcggttgctaggtgcttattggatgcattcctcaagtcccgccctccatctcctat
tcccctctcttctggtttagtgctttaggaaatgtggcagaaatctttttctgcctgtg
tctaggaaatcataattcatgctggcgtaccctggttgttgaggtccctgaatcctTGT
GCCCACACTGCTGAAGACTCCTTGTGTGA[ct]ACAAGTCAGGGGACATCTGGGTCTTG
ACTCCCcagatgctccagctggaccctgctgccctcccttgcccaccctcttccattgt
agatgccaaggggctgagcgatccagggaagatcaagcggctgcgttcccaggtgcagg
tgagcttggaggactacatcaacgaccgccagtatgactcgcgtggccgctttggagag
ctgctgctgctgctgcccaccttgcagagcatcacctggcagatgatcgagcagatcca
gttcatcaagctcttcggcatggccaagattgacaacctgttgcaggagatgctgctgg
gaggtccgtgccaagcccaggaggggcggggttggagtggggactccccaggagacagg
cctcacacagtgagctcacccctcagctccttggcttccccactgtgccgctttgggca
agttgcttaacctgtctgtgc

FIGURE 23tt

47910
TCATTTTTACACAGGATGTACGCGTTTTGAAGCACAAAACTCTCCAGTGATCACAGGTC
ATAGACTGTCTGATTTTTATGTGAAATCCCATTTTAAGAGTAAAATATAAGTAACATAG
TAGGCTCTAGTCTATAAACAAAGACTTCTATTTATAGTTTGTTTGCCCCCTGAGCCCCA
TCTCATCTGCTGGTGGCATGCACATGCTCTTTATTACCAGTGCGAATATAGCTGGGAAA
CTAATGCCACTCACCATACAGGATGGTTAACATGGACACGGGCATGACAAGGAAACCCA
GCAGCATATCAGCTATGGCAAGTGACATCAGGAAATAGTTGGTGGCATTCTGCAGCTTT
TTCTCTAGGGACACTGCCATGATGACGAGTATGTTTCCAGCAATAGTTAGAATAATCAC
TACGGCTGTCAGTAAAGCAGACCAGTTTTTTTCCTGGAGATGAAGTAAGGAGAGACACG
ACGGTGAGAGGCACCCTTCACAGGAAAGGTTGGTTCGATTTTCAGAGTCGACTGTCCAG
TTAAATGCATCAGAAGTGTTAGCTTCTCC [ga] GAGTTAAAGTCATTACTGTAGAGCCT
GGTGTCATCATTTAATTGCATTAGGGAGTTCGTAGTTGAGCTCAAAGAAGTATTTTCTT
CACAAAGAATATCCATGTCTAAGCCAGAACTTGTAGCAGATGAGGTGTAGAAGGACTAA
CAGGTTATAGTTTCTGCTCACCATTCACCTTGATGTACCCACACTCTGTAACACTGAGG
CTGGTGTACATGCTGTTCTCCCGGGGCTGGATTTTTGTCTTCCATTATTACAATGATAG
TTAAAGAACTGAACTGTGGTGGCTGTAAGTTTTCTTCATTCACAATTTTAGGAGAGTCC
ACTGTTTGGTTTTATTATTTTCTCACCAAACCGAGGACAAAAAAGCAGAATGAACTTTT
AGCATAGAGGTTGCAGGGTTTTTTTGAGCGCTCGGGAAGATAAATGTCCTGGACAAAG
AAGAAAAGTTTTATAACTACT

tcatttttacacaggatgtacgcgttttgaagcacaaaactctccagtgatcacaggtc
atagactgtctgatttttatgtgaaatcccattttaagagtaaaatataagtaacatag
taggctctagtctataaacaaagacttctatttatagtttgtttgcccccctgagcccca
tctcatctgctggtggcatgcacatgctctttattaccagtgcgaatatagctgggaaa
ctaatgccactcaccatacaggatggttaacatggacacgggcatgacaaggaaaccca
gcagcatatcagctatggcaagtgacatcaggaaatagttggtggcattctgcagcttt
ttctctagggacactgccatgatgacgagtatgtttccagcaatagttagaataatcac
tacggctgtcagtaaagcagaccagtttttttcctggagatgaagtaaggagagacacg
acggtgagaggcacccttcacaggaaaggttggttCGATTTTCAGAGTCGACTGTCCAG
TTAAATGCATCAGAAGTGTTAGCTTCTCC [ga] GAGTTAAAGTCATTACTGTAGAGCCT
GGTGTCATCATTTAATTGCATTAGGGAgttcgtagttgagctcaaagaagtattttctt
cacaaagaatatccatgtctaagccagaacttgtagcagatgaggtgtagaaggactaa
caggttatagtttctgctcaccattcaccttgatgtacccacactctgtaacactgagg
ctggtgtacatgctgttctcccggggctggatttttgtcttccattattacaatgatag
ttaaagaactgaactgtggtggctgtaagttttcttcattcacaattttaggagagtcc
actgtttggttttattattttctcaccaaaccgaggacaaaaaagcagaatgaactttt
agcatagaggttgcagggtttttttgagcgctcgggaagataaatgtcctggacaaag
aagaaaagttttataactact

FIGURE 23uu

47915
GAAGATTGTGGAAAATGATGGAAGATTCCGGAAAGTGGTGGAAGATTCCAGAAAATGAT
GGAAGATTCCAGAAAGTGATGAAAGATTCTGGAAAGCAATGAAACATTCCAGAAAGTGA
TGAGACAGTGATAGAGTCTGGTTCCAGGCGAAGTGGGAGAGGATGGGATTTGAGAAGGG
AATGATCCCTCCTCACACCTCTAGGATGGGAAGCTTAGTGGAGTGAGGGGTGGGTAGGA
GGTTACACCCTGTGTCCTCTGTCGCTCTGTGCAGGAGGAGGAGGCAGAGAAAGGGAAGG
GTCAGGAAAGCCAGCCCATGTCCCACCCCCACTGGACTCACCACGTGATGGCAGGTGAA
GCCCTTCATGACCGAGGCCTCATTGAGGAACTCAATCCGCTCTCGGAGACTGGCTGACT
CGTTGACCGTCTTCACCGCCACGCGGGTCTCTGCCTCACCCTTGATGATGTCCCTGGCA
TTGCCCTCATACACCATGCCGAAGGAGCCCTGCCCCAGCTCTCGAAGGAGGGTGATCTT
CTCTCGAGACACCTCCCACTCGTCCGGCA[ct]GTACACAGAGCATGGAAACACTACTT
CTTACTTATCTACACAGCATCCTTGGAGGATCCCTTGGGGGTCTGCAGCCACCTTCCAC
CCAAGCCCTCACCCAAACCCCCTCGAAAACACTCATGAAATGAGTTCTGTGATCCAGGA
CCCATGCCGGGCACTGGGCATATGGCCGAGAACAGGACAGGCATCTGCACCCATGGAGA
GGGCATGGCAGAGACTCAAGGAAGGAGCCACAACTGGTCCAAGATCCTGGCCAATATGT
CCTGAGGCAAACCTGCATCCCCATCCTTCTTGTCTGATTTCAGACCCTTGCTATGGAAT
GATGCTACTTCCCACCTGAGACTACTGTTTCTGCAAAGTGCCAAGGGGATGGAAGACAG
GTTGTAATAGGTTGGGGAAAAAAAAAGCCAGGATACTTGGAGCTCTTCCCATGAAAAGG
TGGAGTCTATCTCACCACCCC

gaagattgtggaaaatgatggaagattccggaaagtggtggaagattccagaaaatgat
ggaagattccagaaagtgatgaaagattctggaaagcaatgaaacattccagaaagtga
tgagacagtgatagagtctggttccaggcgaagtgggagaggatgggatttgagaaggg
aatgatccctcctcacacctctaggatgggaagcttagtggagtgaggggtgggtagga
ggttacaccctgtgtcctctgtcgctctgtgcaggaggaggaggcagagaaagggaagg
gtcaggaaagccagcccatgtcccacccccactggactcaccacgtgatggcaggtgaa
gcccttcatgaccgaggcctcattgaggaactcaatccgctctcggagactggctgact
cgttgaccgtcttcaccgccacgcgggtctctgcctcacccttgatgatgtccctggca
ttgccctcatacaccatgccgaaggagccctgccccagCTCTCGAAGGAGGGTGATCTT
CTCTCGAGACACCTCCCACTCGTCCGGCA[ct]GTACACAGAGCATGGAAACACTACTT
CTTACTTATCTACACAGCATCCTTggaggatcccttgggggtctgcagccaccttccac
ccaagccctcacccaaaccccctcgaaaacactcatgaaatgagttctgtgatccagga
cccatgccgggcactgggcatatggccgagaacaggacaggcatctgcacccatggaga
gggcatggcagagactcaaggaaggagccacaactggtccaagatcctggccaatatgt
cctgaggcaaacctgcatccccatccttcttgtctgatttcagacccttgctatggaat
gatgctacttcccacctgagactactgtttctgcaaagtgccaaggggatggaagacag
gttgtaataggttggggaaaaaaaaagccaggatacttggagctcttcccatgaaaagg
tggagtctatctcaccacccc

FIGURE 23vv

47917
TGTATTTTTGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGGC
CTCAAGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTTGGATTACAGGTGTGAGCCACT
GTACCCAGCAATTTATAAGGTTTTAAGACTCAAATAACTCCTTCTAAAGTGAAATGAGT
CTCCTGTTGTGGTGGGAGGCAGACATCATTCAACTTAGAGGACACAGCTGGAAAGCAAT
GTGAGAAACTAAGAAAAGTAACAAGCTGGTAGATTGGCATTTCTGACCCATCTTCCTGC
GAAGTCAGGTATCAAGGCTTTAAGTACTAATAGCACAGTACCTGATGAGAGAAGCACTG
GAATCAAAATTTCAGCAGAGGAAGGAGGTACCAAGTGCAACTCTGAAGGGGCATGCTGA
AGTGTGCAGGGGCATGCCCAAGAGTCAAGGGCCTTACCTCATCACCATATCGCCGATAA
CTCACTTCATACAGCACGATCAGACCATTGGGCTCCTTCGGCTCCTGCCACATCAAGTG
GACGACGTTGTTCTCAAAGATTTCATGCG[tc]CACAGGGCCAACAATGTCATCAGCCT
TGGCTGTAAGGAGAGGAAGTGAGAGGCAGGGATGTAACTCTTGGATGAGATCCCACTTC
TGCCACCTGTCCATGGTGCAACCTTGGGCTGGTGACGTCATTTTCCCACAACCCATTTT
CCTCGTCAGAGAACGGACATCTAAAACTCATCCCACAAGATTGTTAGGAAGATTAAATG
GGTTACTTTCTGCGTATAACTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCACCC
AGGCGGGAGTGCAGTGGTGTATTTTCTAAAGTTTACATAATGATTGCCTATGACTCATA
ATTTTAAAATATGACCTGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGCT
CAAGGTTGGCGGACCACTTGAGCTCAGGCATTGGAGACCAGCCTGGGCAACATGGTGGA
ACCATCTCTACTGAAAATACA

tgtatttttgtagagatggggtttcaccatgttggccaggctggtctcaaactcctggc
ctcaagtgatctgcctgcctcggcctcccaaagtgcttggattacaggtgtgagccact
gtacccagcaatttataaggttttaagactcaaataactccttctaaagtgaaatgagt
ctcctgttgtggtgggaggcagacatcattcaacttagaggacacagctggaaagcaat
gtgagaaactaagaaaagtaacaagctggtagattggcatttctgacccatcttcctgc
gaagtcaggtatcaaggctttaagtactaatagcacagtacctgatgagagaagcactg
gaatcaaaatttcagcagaggaaggaggtaccaagtgcaactctgaaggggcatgctga
agtgtgcaggggcatgcccaagagtcaagggccttacctcatcaccatatcgccgataa
ctcacttcatacagcacgatcagaccattgggctccttcggctcctgccacatcaagtG
GACGACGTTGTTCTCAAAGATTTCATGCG[tc]CACAGGGCCAACAATGTCATCAGCCT
TGGCtgtaaggagaggaagtgagaggcagggatgtaactcttggatgagatcccacttc
tgccacctgtccatggtgcaaccttgggctggtgacgtcattttcccacaacccatttt
cctcgtcagagaacggacatctaaaactcatcccacaagattgttaggaagattaaatg
ggttactttctgcgtataactttttttttttttgagacagagtcttgctctgtcaccc
aggcgggagtgcagtggtgtattttctaaagtttacataatgattgcctatgactcata
attttaaaatatgacctggcatggtggctcatgcctgtaatcccagcactttgggagct
caaggttggcggaccacttgagctcaggcattggagaccagcctgggcaacatggtgga
accatctctactgaaaataca

FIGURE 23ww

47920
GACTGAGGTTCACCCGGGTGAAGGCGCTCATGCCCCCAGGTCCTTGTGGCCCCCCAGC
AGGGACGAGTGGGCAGCCAGCTCTGCTGCCCCTTGAGGCCCAGTCGGGGAAGCAGAGGC
TGCTGAGGATGAGGAGGCAGCAGCCATGGTGGCCCTGGGCAGGCTCACCTCCTCTGCAG
CAATGCCTGTTCGCATGTCAGCATAGCTTACAGGGGCAGCTGGCGAGGTGTCCACGTAG
CTCTGACGGGGACAACTCATCTGCATGGTCATGTAGTCACCCCGGCTGCTGGGCACTGC
CCGGGTAGGCCTGCAAATGCTAGCAGCCCCGGGAGGTGCAGGGCCCAGTCTGCCCATCT
CGACCCCAGTGCTCTCCTGCCAGGCTGCCCTCCGCCCGGCCCCAGGTCCATCTTCATGT
ACTCCTCAGTGCCAGTCTCTTCCTCTCTGGGAGCTGGCTGGAGCTGGGATGGACACCTG
ACAGAAGGTGAGCTGTGGAAAGCCACCGGGCCAGACAAGTAGCCAGACTGATCACTCCC
AAATTCAATATTGACATATTCCCCCGGGC[tc]CTTGGGCTCTGGAGGGTGCAGCAAGG
GCTGCTGCTGCTGCTGCTGCTCTCGGGCCCGAGGTAAGGTGCTGGCCTTGGGATCCCCC
AGGGACAGCCTCGTGGGCCGGGCCAGGCGGCTATTGGTCTGAGCAGCTGTGTCCACCTT
TCGAGGCAGATGGGGCTGCAGAACCTGATGGTGGGGATGTGGAAGGCTGGGCTCCAGCC
TAGCCCCGCAGTATCCCCCACCCAGGCTGTCGCTGCTGGTGGAAGAGGAAGAATCATCT
GCTGTTGCAGCATAGAGAAGGCGACCAGAGCTAGTGGAAAGGCGGAGGTGCTGATGCCG
GGCACCCTCCTCCGGCTCCCCGGGGCGCTGGGTGTGCTTAAAGGATCTTGGCAATGAGT
AGTAGGAGAGGACTGGCTTGTGCTGGGGGTCCTCAGGGCCGTAGTAGCAGTCGGAGGGG
CTGCTGGTGTTGGAGTCCCCC

gactgaggttcacccgggtgaaggcgctcatgcccccaggtccttgtggccccccagc
agggacgagtgggcagccagctctgctgccccttgaggcccagtcggggaagcagaggc
tgctgaggatgaggaggcagcagccatggtggccctgggcaggctcacctcctctgcag
caatgcctgttcgcatgtcagcatagcttacaggggcagctggcgaggtgtccacgtag
ctctgacggggacaactcatctgcatggtcatgtagtcaccccggctgctgggcactgc
ccgggtaggcctgcaaatgctagcagccccgggaggtgcagggcccagtctgcccatct
cgaccccagtgctctcctgccaggctgccctccgcccggccccaggtccatcttcatgt
actcctcagtgccagtctcttcctctctgggagctggctggagctgggatggacacctg
acagaaggtgagctgtggaaagccaccgggccagacaagtagccagactgatcactccc
aaaTTCAATATTGACATATTCCCCCGGGC[tc]CTTGGGCTCTGGAGGGTGCAGCAAGG
gctgctgctgctgctgctgctctcgggcccgaggtaaggtgctggccttgggatccccc
agggacagcctcgtgggccgggccaggcggctattggtctgagcagctgtgtccacctt
tcgaggcagatggggctgcagaacctgatggtggggatgtggaaggctgggctccagcc
tagccccgcagtatcccccacccaggctgtcgctgctggtggaagaggaagaatcatct
gctgttgcagcatagagaaggcgaccagagctagtggaaaggcggaggtgctgatgccg
ggcaccctcctccggctccccggggcgctgggtgtgcttaaaggatcttggcaatgagt
agtaggagaggactggcttgtgctgggggtcctcagggccgtagtagcagtcggagggg
ctgctggtgttggagtccccc

FIGURE 23xx

47922
GATTGGGGATCTGGTGGAAGCGGATGAACTCCCGCACCCGCAGCATCTGTGTGTGGTAG
CGGGCTGTGCCCGAGTACAGCCGCTGGATGATGGCCGACACGTTGCCGAAGATGCTAGC
ATACATGAGGGCTGGGGGCGTGGGCACGTGGGGCCGTCAGCCTCTGCAGGGACCCCACC
CACCCACAGGGACCCTGCTCAGGCCCCGCACCAGGTCAGTGTCTCAGTCTCAGCGTCGA
CATGCCCACGAGACGCCCTTGTACATCTGCGCTCCAGCACACCCCACCCTTCAGTAGTC
CCCGCCCTGGTGACCCAGCCCCCAAACCATGTCACGATGGTGGCCCCTGGAGTCTCTAA
GTTCCAGGGCCTCACTCTGGCCCGGCTAGCAGCCTCAGTTTCCTCCAACTTGGGTTCCT
CCACCGTGGGCTCTCCCCGCCGCCCGCCCCTGGGCACACTCACAGCCAATGAGCATGAC
GCAGATGGAGAAGATCTTCTCTGAGTTGGTGTTGGGAGAGACGTTGCCGAAGCCCACAC
TGGTGAGGCTGCTGAAGGTGAAGTAGAGC[ga]CCGTCACATACTTGTCCTTGATGGAG
GGGCCGCCCAGGCCGCTGCTGTTGTAGGGTTTGCCTATCTGGTCGCCCAGGTTGTGCAG
CCAGCCGATGCGTGAGTCCATGTGTGGCTGCTCCATGTTGCCGATGGCGTACCAGATGC
AGGCTAGCCAGTGCGCGATGAGCGCAAAGGTGCACATGAGCAAGAACAGCACGGCCGCG
CCGTACTCTGAGTAGCGATCCAGCTTCCGCGCCACGCGCACCAGCCGCAGCAGCCGCGC
AGTCTTCAGCAGCCCGATCAGCTGGGGGACAGGGAAGGGGCACATTCCGTTGATGGGGC
AAGGGGGGCAAGGGAGGAGGGGAGGTGCTGCGGCCCTCAGAGCGAGCATCAGAGGTCAG
ATCCCCAAAGACTTCCTAGACCCTCCTCCTAAGAGGTGAAGCCCACACTGGGCCCAGCA
CAGGTGTCTCATTAATCTTAG

gattggggatctggtggaagcggatgaactcccgcacccgcagcatctgtgtgtggtag
cgggctgtgcccgagtacagccgctggatgatggccgacacgttgccgaagatgctagc
atacatgagggctgggggcgtgggcacgtggggccgtcagcctctgcagggaccccacc
cacccacagggaccctgctcaggccccgcaccaggtcagtgtctcagtctcagcgtcga
catgcccacgagacgcccttgtacatctgcgctccagcacaccccacccttcagtagtc
cccgccctggtgacccagcccccaaaccatgtcacgatggtggcccctggagtctctaa
gttccagggcctcactctggcccggctagcagcctcagtttcctccaacttgggttcct
ccaccgtgggctctccccgccgcccgcccctgggcacactcacagccaatgagcatgac
gcagatggagaagatcttctctgagttggtgttgggagagacgttgccgaagcccacAC
TGGTGAGGCTGCTGAAGGTGAAGTAGAGC[ga]CCGTCACATACTTGTCCTTGATGGAG
GGGCCgcccaggccgctgctgttgtagggtttgcctatctggtcgcccaggttgtgcag
ccagccgatgcgtgagtccatgtgtggctgctccatgttgccgatggcgtaccagatgc
aggctagccagtgcgcgatgagcgcaaaggtgcacatgagcaagaacagcacggccgcg
ccgtactctgagtagcgatccagcttccgcgccacgcgcaccagccgcagcagccgcgc
agtcttcagcagcccgatcagctgggggacagggaaggggcacattccgttgatggggc
aaggggggcaagggaggaggggaggtgctgcggccctcagagcgagcatcagaggtcag
atccccaaagacttcctagaccctcctcctaagaggtgaagcccacactgggcccagca
caggtgtctcattaatcttag

FIGURE 23yy

47924
CATGCTCTTGCGGAGGTCACCCACACGTAGCATGAAGCAGAGGCGGCCGTGGCGCAGGG
CGATCACCGCATGCTTGCTGAAGATGAGGGTCTCAGCCCTGCGGTGGGCTTGGGCAGTC
TTCATGAAGATGCAGCCAAGCATGATGGCGTTGATCATGAGCCCCACGATGTTCTGCAC
GATGAGGATCAGGATGGCCAGTGGGCACTCCTCAGTCACCATGCGCCCCCCAAAGCCAA
TAGTCACTTGGACCTCAATGGAGAAAAGGAAGGCAGACGAGAAGGAGTGGATGCTGGTG
ACACAGGGCTCAGCAGTGCCCTCGCTGGGGGCCAGGTCACCGTGGGCGAAGGCGATGAG
CCACCAGGCCATGGCGAAGAGCAGCCAGCTGCACAGGAAGGACATGGTGAAGATGAGCA
ATGTGTGTGGCCACTTGAGGTCCACCAGCGTGGTGAACACGTCCTGCAGGAAGCGGCCC
TGCTCCCGGATGTTCTTGTGGGCCACGTTGCAGTTGCCTTTCTTGGACACAAAGCGGGC
CCTCCGCTGGCGGGCACGGTACCTGGGCT[tc]GGCAGGGTCCTCTGCCAGGCGTGTCA
GCACGTATTCCTCGGGGATGATGCCCTTGCGGGACAGCATGGCTCCGGTGACCCCCAGG
GAGGGGCTTCCCCCATCGGAGGCACCCCTCGGACGTGGCCTAGGGCCTCACTGCAGAGT
CCTCTCGGTGGGCACCTTCTCACCCTGGGGCTGCACTCAGCCTGTGCTGGCCTCACTTC
TGAGATAACTCCCCACCAGACTCTTCCTTACCTCCACCTGGGTCCCACTTCACTTCTTA
ATACCAGCCTCAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGTACGTTGGGAGGC
TGAGGAGGGCAGATCACTAGGTCAGGAGTTCGAGACCAGCCTGACCAACATGGTGAAAC
CCCATCTCTACTAAAAATACAAAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAATCCC
AGCTACTCAGGAAGCTGAGGC

catgctcttgcggaggtcacccacacgtagcatgaagcagaggcggccgtggcgcaggg
cgatcaccgcatgcttgctgaagatgagggtctcagccctgcggtgggcttgggcagtc
ttcatgaagatgcagccaagcatgatggcgttgatcatgagccccacgatgttctgcac
gatgaggatcaggatggccagtgggcactcctcagtcaccatgcgcccccaaagccaa
tagtcacttggacctcaatggagaaaaggaaggcagacgagaaggagtggatgctggtg
acacagggctcagcagtgccctcgctgggggccaggtcaccgtgggcgaaggcgatgag
ccaccaggccatggcgaagagcagccagctgcacaggaaggacatggtgaagatgagca
atgtgtgtggccacttgaggtccaccagcgtggtgaacacgtcctgcaggaagcggccc
tgctcccggatgttcttgtgggccacgttgcagttgcctttcttggacacaaagcgggc
cctccgcTGGCGGGCACGGTACCTGGGCT[tc]GGCAGGGTCCTCTGCCAGGCGTgtca
gcacgtattcctcggggatgatgcccttgcgggacagcatggctccggtgacccccagg
gaggggcttcccccatcggaggcacccctcggacgtggcctagggcctcactgcagagt
cctctcggtgggcaccttctcaccctggggctgcactcagcctgtgctggcctcacttc
tgagataactccccaccagactcttccttacctccacctgggtcccacttcacttctta
ataccagcctcaggccgggcgcggtggctcacgcctgtaatcccagtacgttgggaggc
tgaggagggcagatcactaggtcaggagttcgagaccagcctgaccaacatggtgaaac
cccatctctactaaaaatacaaaagttagccgggcatggtggtgcgcacctgtaatccc
agctactcaggaagctgaggc

FIGURE 23zz

47925
CACTTCTTGGAGCCACAGACGCAAAGCAGCAGCCCTCGGGGATTGTTCTTCCCCAGCCA
CCGGCCCAGAGTGTGGCTGGTCAATCGTGGGGACCCAGGACTGGCTGGACGCACAGCTC
TAGGGCCCAGTACCTCCCACAGCCTCTGCAGCCTTGGGCGGGGGAGAGGGGTGAGCCAG
TCCTGAATTGGGTTGGGAGGAGCAGGGACAAAAATAACCCAGTACAGGTTCCTGCTGAG
GCCAGAAATAGCATAGTGACAAGTGCCTTGTAACACCCTGGATGAGCAGCAGGGGGAGG
CTGAGCTGAGGCTGGCCCAGCCTCACACCAGGCCCTGGCCGGGCTACATACCACATGGT
CCGTGTGTACACACGCGTGTGGGGGGCCCGAGAGACCATGGCTCAGGACAGGGAATCTG
GAGAGATGCTGAACTTGGGCTTGGCCTTGGCCATGGGCACGCTGCGCTTGCGCAGGGGC
CCGCGGGCTGAGGCGAGGGTCAGAGCTTCCAGTAGGCTGTGGTCCTCATCAAGCTGGCG
GGCCGTGCAGAGTGGTGTGGGCACTTTGA[ct]GGTGTTGCCAAACTTGGAGTAGTCCA
CAGAGTAACGTCCGTCCTCCTCAGCTACAATGGGCACAAAGCGCTGGCCCCACAGGATC
TCATCGGCCAGGTAGGAGGTGCGGGCCTGGGTGGTGATGCCCGTGGTTTCCACCACGCC
TTCCAGGATGACGATGATCTCGAGGTCCTGGTGGTGGTGCAGGTCGCTGGGTGCCAGGT
CGTAGAGTGGGCTGTTGGCATCAATGACATGGTAGATGATCAGCGGGGCCACCAGGAAG
ATGCTGTTGCCACCCACGCCGTTCTCCATGGGGATGTCCACCTGGTGGAGGGGCACCAC
CTCGCCCTCGGGGCTGGTGGTCTTGCGTACCACCTGCATGTGGATGGTGGCGCTGATGA
TCATGCTCTTGCGGAGGTCACCCACACGTAGCATGAAGCAGAGGCGGCCGTGGCGCAGG
GCGATCACCGCATGCTTGCTG

cacttcttggagccacagacgcaaagcagcagccctcggggattgttcttccccagcca
ccggcccagagtgtggctggtcaatcgtggggacccaggactggctggacgcacagctc
tagggcccagtacctcccacagcctctgcagccttgggcgggggagaggggtgagccag
tcctgaattgggttgggaggagcagggacaaaaataacccagtacaggttcctgctgag
gccagaaatagcatagtgacaagtgccttgtaacaccctggatgagcagcagggggagg
ctgagctgaggctggcccagcctcacaccaggccctggccgggctacataccacatggt
ccgtgtgtacacacgcgtgtggggggcccgagagaccatggctcaggacagggaatctg
gagagatgctgaacttgggcttggccttggccatgggcacgctgcgcttgcgcaggggc
ccgcgggctgaggcgagggtcagagcttccagtaggctgtggtcctcatcaAGCTGGCG
GGCCGTGCAGAGTGGTGTGGGCACTTTGA[ct]GGTGTTGCCAAACTTGGAGTAGTCCA
CAGAGTAACGTccgtcctcctcagctacaatgggcacaaagcgctggccccacaggatc
tcatcggccaggtaggaggtgcgggcctgggtggtgatgcccgtggtttccaccacgcc
ttccaggatgacgatgatctcgaggtcctggtggtggtgcaggtcgctgggtgccaggt
cgtagagtgggctgttggcatcaatgacatggtagatgatcagcggggccaccaggaag
atgctgttgccacccacgccgttctccatggggatgtccacctggtggaggggcaccac
ctcgccctcggggctggtggtcttgcgtaccacctgcatgtggatggtggcgctgatga
tcatgctcttgcggaggtcacccacacgtagcatgaagcagaggcggccgtggcgcagg
gcgatcaccgcatgcttgctg

FIGURE 23aaa

47927
ATCTGGTATTGTACAACACATGCAGGTAAGTAACTGAAATCTCCAGGAGTTGGATGTGT
AGTATTTTGGGAGGAGACCAGGCTTGGGCCACAAATGAGGGCACTTTGCACTTTCATCA
AATCCATGTCTACCTTGTCAATCTGAATAACTGAGAGAGGGCAGGTAGATATTTTACAC
CTTGAAGATTTGTTTTCTGGTCATGTAAAAATTAAATATAAACAAATAAAGAACAAAGC
AAGAGAGACAGAAAAAGAAAGAGAATGAGAGACAAGGAAAGATTGTGTTGGGGGGAGAA
GAGAAGGGTTTGCCCAGCTAGGGCACTAAACTTTGGATTCATTCTCCAGGTTTGCCACA
TCACCATTTCTTTCTGTTTGCTCTTCGAGGTTCTTTTCTTCCTCTTCAGTCTCCAGTTC
TGCATGTTGGTTGAGTTTGCTGGATACAGACCAACTCAGGGGCAGCTCTGCCCTGCTGG
CTAACTCGGCCAGCTCTTTGGCACTAAGGGATGGGGTGCTGGTCTCATAGGTCTCATGG
AAGCTGTTGTAGTCAACTTCGTAGAACCC[ga]TCCTCCAGGGTCAGGACAGGTGTGAA
CCGGTAACCCCACAGGATCTCACTGGTGATGTAGGAGCTTCGAGCTTGGCATGTCATCC
CTGCAGAGAGAAGAATGGAGGCTTTAGCATATGTAAGTGTGGGCTTTCCATGGCCAAGG
AGTCACAGAGAGCCAGGAGGAGTACTGCATGCAGCTGTTGAGACTGACCTGCATACGAT
GCCACACTTAGTAGGTGTCATTCATGTTGTAGACACATGCTAATGTGCCATGGAGATTC
CAGGCCTCTTAAGGGAGTCCTGGGGAACAATGAGAGAGTCCTGGCCCACATCAAGCCAC
ATTTGCCTGCATGGCCATGCACATGCAAAGGAAATCAAGTGTGCAAATGCACACAAGTT
TTCGCATGTGCATGGCTATGTCTGGTCCACTCTGCTCTGGGAGAACCCTGAAGCCATGA
CTCTGGCCTCCTACTGCTCTT

atctggtattgtacaacacatgcaggtaagtaactgaaatctccaggagttggatgtgt
agtattttgggaggagaccaggcttgggccacaaatgagggcactttgcactttcatca
aatccatgtctaccttgtcaatctgaataactgagagagggcaggtagatattttacac
cttgaagatttgttttctggtcatgtaaaaattaaatataaacaaataaagaacaaagc
aagagagacagaaaaagaaagagaatgagagacaaggaaagattgtgttggggggagaa
gagaagggtttgcccagctagggcactaaactttggattcattctccaggtttgccaca
tcaccatttctttctgtttgctcttcgaggttcttttcttcctcttcagtctccagttc
tgcatgttggttgagtttgctggatacagaccaactcaggggcagctctgccctgctgg
ctaactcggccagctctttggcactaagggatggggtgctggtctcataggtctcatgg
aAGCTGTTGTAGTCAACTTCGTAGAACCC[ga]TCCTCCAGGGTCAGGACAGGTGTGAA
CCggtaaccccacaggatctcactggtgatgtaggagcttcgagcttggcatgtcatcc
ctgcagagagaagaatggaggctttagcatatgtaagtgtgggctttccatggccaagg
agtcacagagagccaggaggagtactgcatgcagctgttgagactgacctgcatacgat
gccacacttagtaggtgtcattcatgttgtagacacatgctaatgtgccatggagattc
caggcctcttaagggagtcctggggaacaatgagagagtcctggcccacatcaagccac
atttgcctgcatggccatgcacatgcaaaggaaatcaagtgtgcaaatgcacacaagtt
ttcgcatgtgcatggctatgtctggtccactctgctctgggagaaccctgaagccatga
ctctggcctcctactgctctt

FIGURE 23bbb

47932
ACGGGTGCCGGTCAAGAGAGGGGGGCACCCCGTGCCTCCCTACCACACCTTCTGGAAGA
CATAGCCCCCGCTGGGGCCCCAGCCCACGATGGGGTCGGAGGACGGCTTCCCGTTGATG
TTGGGCTGTGAGTTGATGGTGAGGATGCCCTGGCGGTTCACCCGCAGCAGCTCCTCCTT
CAGCAGGCTGGTCTCAGCCGCCAGGGGCTCATCGTTCCAGGGCAGGCAAGTCACCTGGG
AGAGACGGTGAGCTGGCTGGGGCGACCATCAGGTTTGGCACCCTGAGTCCCTCTCACGG
CCCCCAACAAAGACCCAGCCTGTCTTTGCCTCCCTAAGCCCTTCCAGGTGGAGGTCTCC
CAACTTACCCTTCTCCCTTTGCCATGTCCACAGCATGGAGGGGAGGGCACAGGATGGGG
AAGTCACAGCCCCGCAGCCTGGCCTGCAGCTGGGGTCAGGCCAGGGGCAGGGGATGAAC
CAGGGTCCCCACTCCAGCATCACTCACTTTGTGACCATTCCGGTTTGGTTCTCCCGAGA
GGTAAAGAACAAAGACTTCAAAGACACTT [tg] CTTCACTGGTCAGCTCCTCCCCCCAC
ATCTTCAGCAGCTCCTCCTTGGGGGACTTGCTCTTCAGGTAGAAGAGGTAGTAGTCCTT
CAGCTCCCCAAAGGCAGGGGAAGAGGAATTGCCCCTGGCAGAGGGGTGCCCAGAGGTCA
GGGCACACTCCTGACAGAGGGCAGTGCCACCACATGCCCAGGAGGCCATTCCTGTAAAT
TCTGCCCCTGACTCCTCCCAGGTCAACCACAAGCATGCAAACTTCTTCTGCCCTCCCGC
TCCCAAGAACAAAGATGTATTTGCAAGGAAGGTCTGCAGGCCCTCACCAGCGGCCGTTA
GGGAACTCGTCCCACTCCTGGGTACGGTAGATGTAACTCTTTGGTCTGGAGGCCCAGAA
GATGGGACGTACATCTTCCTCTCGGCGCTTGGGGTGGGCGCTGAGAGCCCAGGGTAGGG
GACGCCTGGGTGAGGATGGGG

acgggtgccggtcaagagaggggggcaccccgtgcctccctaccacaccttctggaaga
catagcccccgctggggccccagcccacgatggggtcggaggacggcttcccgttgatg
ttgggctgtgagttgatggtgaggatgccctggcggttcacccgcagcagctcctcctt
cagcaggctggtctcagccgccaggggctcatcgttccagggcaggcaagtcacctggg
agagacggtgagctggctggggcgaccatcaggtttggcaccctgagtccctctcacgg
cccccaacaaagacccagcctgtctttgcctccctaagcccttccaggtggaggtctcc
caacttacccttctccctttgccatgtccacagcatggaggggagggcacaggatgggg
aagtcacagccccgcagcctggcctgcagctggggtcaggccaggggcaggggatgaac
cagggtccccactccagcatcactcactttgtgaccattccggtttggttctcccgaga
ggTAAAGAACAAAGACTTCAAAGACACTT [tg] CTTCACTGGTCAGCTCCTCCCCCCAC
Atcttcagcagctcctccttgggggacttgctcttcaggtagaagaggtagtagtcctt
cagctccccaaaggcaggggaagaggaattgcccctggcagaggggtgcccagaggtca
gggcacactcctgacagagggcagtgccaccacatgcccaggaggccattcctgtaaat
tctgcccctgactcctcccaggtcaaccacaagcatgcaaacttcttctgccctcccgc
tcccaagaacaaagatgtatttgcaaggaaggtctgcaggccctcaccagcggccgtta
gggaactcgtcccactcctgggtacggtagatgtaactctttggtctggaggcccagaa
gatgggacgtacatcttcctctcggcgcttggggtgggcgctgagagcccagggtaggg
gacgcctgggtgaggatgggg

FIGURE 23ccc

47937
GCCACCTCCCTGGATTCTTGGGCTCCAAATCTCTTTGGAGCAATTCTGGCCCAGGGAGC
AATTCTCTTTCCCCTTCCCCACCGCAGTCGTCACCCCGAGGTGATCTCTGCTGTCAGCG
TTGATCCCCTGAAGCTAGGCAGACCAGAAGTAACAGAGAAGAAACTTTTCTTCCCAGAC
AAGAGTTTGGGCAAGAAGGGAGAAAAGTGACCCAGCAGGAAGAACTTCCAATTCGGTTT
TGAATGCTAAACTGGCGGGGCCCCCACCTTGCACTCTCGCCGCGCGCTTCTTGGTCCCT
GAGACTTCGAACGAAGTTGCGCGAAGTTTTCAGGTGGAGCAGAGGGGCAGGTCCCGACC
GGACGGCGCCCGGAGCCCGCAAGGTGGTGCTAGCCACTCCTGGGTTCTCTCTGCGGGAC
TGGGACGAGAGCGGATTGGGGGTCGCGTGTGGTAGCAGGAGGAGGAGCGCGGGGGGCAG
AGGAGGGAGGTGCTGCGCGTGGGTGCTCTGAATCCCCAAGCCCGTCCGTTGAGCCTTCT
GTGCCTGCAGATGCTAGGTAACAAGCGAC[tc]GGGGCTGTCCGGACTGACCCTCGCCC
TGTCCCTGCTCGTGTGCCTGGGTGCGCTGGCCGAGGCGTACCCCTCCAAGCCGGACAAC
CCGGGCGAGGACGCACCAGCGGAGGACATGGCCAGATACTACTCGGCGCTGCGACACTA
CATCAACCTCATCACCAGGCAGAGGTGGGTGGACCGCGGGACCGATTCCGGGAGCGCC
AGTGCCTGCACACCAGGAGATCCTGGGGATGTTAGGGAAAGGGATTGTTTCTTTTCCTT
CGCTCTATCCCAGGGCAGGACAGTATCAGGCACTTAGTCAGCTCTAGGTAAATGTTTGT
ACAGGGCACACTCTACACAAAATGGGTACCTTCCATTTTGTGCAACTACAGTCACAGAG
TCGTGATCCCCAGATTCAGGTTCCCCAGGCTGGTAGGCTGGCAATCTCCTCTCACTCAC
CTCTTATGGTTTGTTGTGGTT

gccacctccctggattcttgggctccaaatctctttggagcaattctggcccagggagc
aattctctttccccttccccaccgcagtcgtcaccccgaggtgatctctgctgtcagcg
ttgatcccctgaagctaggcagaccagaagtaacagagaagaaacttttcttcccagac
aagagtttgggcaagaagggagaaaagtgacccagcaggaagaacttccaattcggttt
tgaatgctaaactggcggggcccccaccttgcactctcgccgcgcgcttcttggtccct
gagacttcgaacgaagttgcgcgaagttttcaggtggagcagaggggcaggtcccgacc
ggacggcgcccggagcccgcaaggtggtgctagccactcctgggttctctctgcgggac
tgggacgagagcggattgggggtcgcgtgtggtagcaggaggaggagcgcggggggcag
aggagggaggtgctgcgcgtgggtgctctgaatccccaagcccgtccgttgagccttct
gtgcctGCAGATGCTAGGTAACAAGCGAC[tc]GGGGCTGTCCGGACTGACCCTCGccc
tgtccctgctcgtgtgcctgggtgcgctggccgaggcgtacccctccaagccggacaac
ccgggcgaggacgcaccagcggaggacatggccagatactactcggcgctgcgacacta
catcaacctcatcaccaggcagaggtgggtggaccgcgggaccgattccgggagcgcc
agtgcctgcacaccaggagatcctggggatgttagggaaagggattgtttcttttcctt
cgctctatcccagggcaggacagtatcaggcacttagtcagctctaggtaaatgtttgt
acagggcacactctacacaaaatgggtaccttccattttgtgcaactacagtcacagag
tcgtgatccccagattcaggttccccaggctggtaggctggcaatctcctctcactcac
ctcttatggtttgttgtggtt

FIGURE 23ddd

47941
ACCCAGAATCCTGCAGTTTCTCCTGATTAACAGCTAAGTAAATTCTATAGCACTGTACT
GAAAATATAAAAAATTTAGAATATAGGGCTGATCATCCCTGATCCTAAGATTGTCCTCT
GAAGTTGATTTTCAGGGTAAATCTTTCATATCCACTTTTTAAATTGCCGATTGTTTCTT
ATGAAACAAGTAGTAAAATGTACAAAAGAAAAAGAATCTAGCTTAAATTATAGAGTTCA
GACATATTTTTTAGTAGGAGGAAGAGGAATAGAATAACAAAATAGAGTGTGAAATTTGG
AGTAAATTGACAGATTTTCAGAATAAAATGTTTCTTTTTTCTCTGTACATGTTAAAAAT
ATACTTTGTATTGATACTTTCATGTGCCATCACTAATATTACATATATAGCATATTAAA
GAGTGACATTTTAAACCATTGTTAAATTATTCAACAGGGACTAAATAGGAATAGTTTGC
CAACTCCACAGCTGAGGAGAAGCTCAGGAACTTCAGGATTGCTACCTGTTGAACAGTCT
TCAAGGTGGGATCGTAATAATGGCAAAAG[ag]CCTCACCAAGAATTTGGCATTTCAAG
GTAAAATCTGCAGAGCCTTTTAAGAAACTTGAATCAAATGCATCTACTTTGTTTCTGTC
AATAATGTTTCAAATAGTTCTGGAAGCAGAAAGGAATGGTTGAAGTATTTTAGGTATAG
GACAACATGTGTAGTAATAATATGGTAAAATAGAGAAACTGATTATTAAAGAGAAGCTA
ATGTGTCTTGTCCTAAAACTTTGATAGGCTGGGTACAAAATGTGCTGGATCCCTGAGAA
CATGAGATAGTTTAGGGAAATCAGGATCAACTCAGGACTGGATGCTGGGGAAGTTTTTA
AATCGATAGAAGTGGCCATTACAGGGTTAGCCACCAATCCAATGAATAGTATCCAAAGG
TAGGTCTGCAGAATTACTGACTTCTGAAAAGAGGAGCACGTTTCCAAGGCTCATCACAA
TTGTTAGGTTTAAGGTAACCA

acccagaatcctgcagtttctcctgattaacagctaagtaaattctatagcactgtact
gaaaatataaaaaatttagaatatagggctgatcatccctgatcctaagattgtcctct
gaagttgattttcagggtaaatctttcatatccactttttaaattgccgattgtttctt
atgaaacaagtagtaaaatgtacaaaagaaaaagaatctagcttaaattatagagttca
gacatattttttagtaggaggaagaggaatagaataacaaaatagagtgtgaaatttgg
agtaaattgacagattttcagaataaaatgtttcttttttctctgtacatgttaaaaat
atactttgtattgatactttcatgtgccatcactaatattacatatatagcatattaaa
gagtgacattttaaaccattgttaaattattcaacagggactaaataggaatagtttgc
caactccacagctgaggagaagctcaggaacttcaggattgctaccTGTTGAACAGTCT
TCAAGGTGGGATCGTAATAATGGCAAAAG[ag]CCTCACCAAGAATTTGGCATTTCAAG
GTAAAATCTGCAGAGCcttttaagaaacttgaatcaaatgcatctactttgtttctgtc
aataatgtttcaaatagttctggaagcagaaaggaatggttgaagtattttaggtatag
gacaacatgtgtagtaataatatggtaaaatagagaaactgattattaaagagaagcta
atgtgtcttgtcctaaaactttgataggctgggtacaaaatgtgctggatccctgagaa
catgagatagtttagggaaatcaggatcaactcaggactggatgctggggaagttttta
aatcgatagaagtggccattacagggttagccaccaatccaatgaatagtatccaaagg
taggtctgcagaattactgacttctgaaaagaggagcacgtttccaaggctcatcacaa
ttgttaggtttaaggtaacca

FIGURE 23eee

47950
CTCCTTTAACATAAGATATATGGGTAAGAAAATTCCAATTTAATGATATTCAAATATAT
AAATATTTGTTGCATCCTCAGGTTTCTAGTTATGTGTTAAAAAAATGATATGTTGAAAT
CTCTTCAATTTTAGAAGAACCTTGTTATAAAGAACAGAGCTAAAAATATTAGAACCACC
TGCCCTTTAGTGTAACAAAATAAACTAGCCTTTTTGGTTTACTTAATTACAGTCTTACC
ATCAAAAATATATTCTCTAACTTAAAAAAATACTTTTTTGGTAATATTTGATGACATTT
CTGATGAGAGCACATAAAAATAAAACAATACTTAAAGATGTGGATATAAAATGCTCAAG
GAATCATCATTTAAAAACAGACGGTTCCCTTATTGTTTCTGTTCATGTCAAAAAGCAGG
GTTTTTTTTTTACACAGTCTCTGTAGCTCCTAGGAATTTCATTTCTACAGCAGCTTTTG
GCCTGTGGGCTGAGCCACTCTTCTTTTGGAATTCTGCAGCAATTTCCTCAAAAGACTTT
CCTTTGGTTTCTGGAACTTTAAAAAATGT [ga] AACAGGGTAAAGGCCAGGAGCACTCC
AGCAAAGAGGAAAAACACATAAGGTCCACAGAAGTCCTGGATAGAAAGCAAACACAGAC
TTTGAGTTAGCAGTTTTTTGACCCTCTCTTCTGTTCAGTAAATCTGTGGAATATTAGGC
TGCTTACCGCAATGTACTGGAAACACAGAGCTACAATGAAATTGCAGGTCCAATTGCTG
AATGCAGCTATTGCTAAAGCAGCAGGACGTGGTCCTTGACTGAAAAACTCAGCCACCAT
GAACCAGGGGATCGGGCCTGGCCCAATTTCAAAGAAGCTGACAAAGAGGAAGATGGCTA
TCATGCTCACATAACTCATCCAAGAGAACTTATTCTGAGGAAAAAAACAAAAACAATAG
TGGGACTGAGATCATTTGGCTGCTTTTTCCTTTAGCTAAGTAGCCTCTGAGTTCACAGG
CGGCATACAACTTTTTCTAAT

ctcctttaacataagatatatgggtaagaaaattccaatttaatgatattcaaatatat
aaatatttgttgcatcctcaggtttctagttatgtgttaaaaaaatgatatgttgaaat
ctcttcaattttagaagaaccttgttataaagaacagagctaaaaatattagaaccacc
tgccctttagtgtaacaaaataaactagcctttttggtttacttaattacagtcttacc
atcaaaaatatattctctaacttaaaaaaatacttttttggtaatatttgatgacattt
ctgatgagagcacataaaaataaaacaatacttaaagatgtggatataaaatgctcaag
gaatcatcatttaaaaacagacggttcccttattgtttctgttcatgtcaaaaagcagg
gttttttttttacacagtctctgtagctcctaggaatttcatttctacagcagcttttg
gcctgtgggctgagccactcttcttttggaattctgcagcaatttcctcaaaagacttt
CCTTTGGTTTCTGGAACTTTAAAAAATGT [ga] AACAGGGTAAAGGCCAGGAGCACTCC
AGCaaagaggaaaaacacataaggtccacagaagtcctggatagaaagcaaacacagac
tttgagttagcagttttttgaccctctcttctgttcagtaaatctgtggaatattaggc
tgcttaccgcaatgtactggaaacacagagctacaatgaaattgcaggtccaattgctg
aatgcagctattgctaaagcagcaggacgtggtccttgactgaaaaactcagccaccat
gaaccaggggatcgggcctggcccaatttcaaagaagctgacaaagaggaagatggcta
tcatgctcacataactcatccaagagaacttattctgaggaaaaaaacaaaaacaatag
tgggactgagatcatttggctgctttttcctttagctaagtagcctctgagttcacagg
cggcatacaacttttctaat

FIGURE 23fff

47960
ATGGAATACAGGGGACGTTTAAGAAGATATGGCCACACACTGGGGCCCTGAGAAGTGAG
AGCTTCATGAAAAAAATCAGGGACCCCAGAGTTCCTTGGAAGCCAAGACTGAAACCAGC
ATTATGAGTCTCCGGGTCAGAATGAAAGAAGAAGGCCTGCCCCAGTGGGGTCTGTGAAT
TCCCGGGGGTGATTTCACTCCCCGGGGCTGTCCCAGGCTTGTCCCTGCTACCCCCACCC
AGCCTTTCCTGAGGCCTCAAGCCTGCCACCAAGCCCCCAGCTCCTTCTCCCCGCAGGGA
CCCAAACACAGGCCTCAGGACTCAACACAGCTTTTCCCTCCAACCCCGTTTTCTCTCCC
TCAAGGACTCAGCTTTCTGAAGCCCCTCCCAGTTCTAGTTCTATCTTTTTCCTGCATCC
TGTCTGGAAGTTAGAAGGAAACAGACCACAGACCTGGTCCCCAAAAGAAATGGAGGCAA
TAGGTTTTGAGGGGCATGGGGACGGGGTTCAGCCTCCAGGGTCCTACACACAAATCAGT
CAGTGGCCCAGAAGACCCCCCTCGGAATC[tc]GAGCAGGGAGGATGGGGAGTGTGAGG
GGTATCCTTGATGCTTGTGTGTCCCCAACTTTCCAAATCCCCGCCCCCGCGATGGAGAA
GAAACCGAGACAGAAGGTGCAGGGCCCACTACCGCTTCCTCCAGATGAGCTCATGGGTT
TCTCCACCAAGGAAGTTTTCCGCTGGTTGAATGATTCTTTCCCCGCCCTCCTCTCGCCC
CAGGGACATATAAAGGCAGTTGTTGGCACACCCAGCCAGCAGACGCTCCCTCAGCAAGG
ACAGCAGAGGACCAGCTAAGAGGGAGAGAAGCAACTACAGACCCCCCCTGAAAACAACC
CTCAGACGCCACATCCCCTGACAAGCTGCCAGGCAGGTTCTCTTCCTCTCACATACTGA
CCCACGGCTCCACCCTCTCTCCCCTGGAAAGGACACCATGAGCACTGAAAGCATGATCC
GGGACGTGGAGCTGGCCGAGG

atggaatacaggggacgtttaagaagatatggccacacactggggccctgagaagtgag
agcttcatgaaaaaaatcagggaccccagagttccttggaagccaagactgaaaccagc
attatgagtctccgggtcagaatgaaagaagaaggcctgccccagtggggtctgtgaat
tcccgggggtgatttcactccccggggctgtcccaggcttgtccctgctacccccaccc
agcctttcctgaggcctcaagcctgccaccaagcccccagctccttctccccgcaggga
cccaaacacaggcctcaggactcaacacagcttttccctccaaccccgttttctctccc
tcaaggactcagctttctgaagcccctcccagttctagttctatctttttcctgcatcc
tgtctggaagttagaaggaaacagaccacagacctggtccccaaaagaaatggaggcaa
taggttttgaggggcatggggacggggttcagcctccagggtcctacacacaaatcagt
caGTGGCCCAGAAGACCCCCCTCGGAATC[tc]GAGCAGGGAGGATGGGGAGTGTGAGG
Ggtatccttgatgcttgtgtgtccccaactttccaaatccccgcccccgcgatggagaa
gaaaccgagacagaaggtgcagggcccactaccgcttcctccagatgagctcatgggtt
tctccaccaaggaagttttccgctggttgaatgattctttccccgccctcctctcgccc
cagggacatataaaggcagttgttggcacacccagccagcagacgctccctcagcaagg
acagcagaggaccagctaagagggagagaagcaactacagaccccccctgaaaacaacc
ctcagacgccacatcccctgacaagctgccaggcaggttctcttcctctcacatactga
cccacggctccaccctctctcccctggaaaggacaccatgagcactgaaagcatgatcc
gggacgtggagctggccgagg

FIGURE 23ggg

47963
GAGTCCCCTCCTTACTGGGGTCCCTGCCCCAGCCTGAGGGGAGGGAAAGCTCTGCCTAA
GACCGCCTGCGTCCAGAGTCCAGACCTACCTTTCCACAGGCCCCTGACTCCTTCCTCCC
TGGCGATGGTTCTGTAGGCGTCCATAGTCCCGCTGTATTTTCTGTCGCTCCTGGATGGC
CCGAGGTGTATGCTGGCCTGAAATCGGACCTTCACCACATCTGTGGGCTGGGCACAGGT
CACCGCCATGGCTCCTGTGGTGCAGCCGGCCAAAATCCGGGTAGTGAGGCTGGAGTCTG
GGAGGGGCAGAGAGAGTGGGCCAGTGTCCCCTACTAAGCAGCATTCTGGGACATGCTGT
TCTCTGCGGGGCTGCCCCTGCAGCTTCCTTGATGTCCACTCAGAGCCTCCTCATAAGCG
TCCGGTACCAGCTTCCCCCCGCCCCTGGCTCTGCCTCTGAGTCTAGACTTCCCTGGTCT
CTTGACCCACACACTTTCAGCCACCCCTTTGGTGTTCAGGGACCTGGTCACTCACTGTC
CGCGCCTTTGGGGGTGTACACCTGCTTGA[ct]GGAGTCATAGAGGCCGATGCGGATGG
AGGCGAAGCTCATCTGGCGCTGCAGGCCGGCCACCAGCCCATTGTAGGGGCTGCAGGGA
CCCTCAGTCCGCACCATGGTCAGGATGGTGCCCAGCACGCCACGGTACTGCACGAGCCG
GGCCGTCTGGACCGCCTGGTTCTCCCCCTGGATCTGAGGGACAATAGCAGGGGGTGAGG
ACTCAGATGGGAAGGCAAGAAGGGGCTGCGTGCACAGGAACCCTGCTGGGGCTGGGCCT
GCCTGGGCTGGGCCTGAGAACAACCATGCTGGTCACAGTAGAAATCACTGGTGTCTGCG
CAGCATTTTACCATTCACAAAGCAGTATTATACACATGGCTTGGTGTTTGATCCTCAGA
GTAAATCAGAGGGACAGATTGTTTTTCCCATTTTATAAGTGCTTCGTGGCTTGCCCAAG
GTCACACAGTTAATTCCTTAC

gagtcccctccttactggggtccctgccccagcctgaggggagggaaagctctgcctaa
gaccgcctgcgtccagagtccagacctacctttccacaggcccctgactccttcctccc
tggcgatggttctgtaggcgtccatagtcccgctgtattttctgtcgctcctggatggc
ccgaggtgtatgctggcctgaaatcggaccttcaccacatctgtgggctgggcacaggt
caccgccatggctcctgtggtgcagccggccaaaatccgggtagtgaggctggagtctg
ggaggggcagagagagtgggccagtgtcccctactaagcagcattctgggacatgctgt
tctctgcggggctgcccctgcagcttccttgatgtccactcagagcctcctcataagcg
tccggtaccagcttccccccgcccctggctctgcctctgagtctagacttccctggtct
cttgacccacacactttcagccacccctttggtgttcagggacctggtcactcactgtc
CGCGCCTTTGGGGGTGTACACCTGCTTGA[ct]GGAGTCATAGAGGCCGATGCGGATGG
AGGcgaagctcatctggcgctgcaggccggccaccagcccattgtaggggctgcaggga
ccctcagtccgcaccatggtcaggatggtgcccagcacgccacggtactgcacgagccg
ggccgtctggaccgcctggttctccccctggatctgagggacaatagcagggggtgagg
actcagatgggaaggcaagaaggggctgcgtgcacaggaaccctgctggggctgggcct
gcctgggctgggcctgagaacaaccatgctggtcacagtagaaatcactggtgtctgcg
cagcattttaccattcacaaagcagtattatacacatggcttggtgtttgatcctcaga
gtaaatcagagggacagattgtttttcccattttataagtgcttcgtggcttgcccaag
gtcacacagttaattccttac

FIGURE 23hhh

47972
AAGAAAATCAAACTTAACTCGGACCCAGAGACATTTTAGTATGTGTTGGAAACTTTAGC
ATCTGGTCACCATCCTCCAAAGAATTATTTGGATTGGAACTCGGTCAGAGCTGTCACTC
TTCAGCTAGGAATCTAAGAGGATCATGTCTTGGATGTTACGGAGTATAGACAACCAAGT
TCCCTGCCCTCAAAAGCCCGATCACTTATAAGACAGCTTATGGAGCTTTGACAGAGGGC
AGCAGTTGATGGCATTATCCTTTGAACTCATAGCTTAGTTGGACTCCTACTGGCTTGTG
GGACCAAATCTTTCCCTACCACAGTTGGCTATAGCAAAAGTTGTGAAAAATGCCACTAG
GATATACTGGTGAGGGAAAAGGAGGTCCATTTGTAGTTATAGTATAATTGAAAAGAAAA
GCTCTGAAGAAAACTCTAGCCTACTCTTTTTCAGCCCAAGGGGAAGGCAGAGCACCTGC
TGACAGATGCTGGCGTAGCGAGCCAGGGCGTTGGCGTTTTCCTGGATAGCGAGGCTGGA
TGGACACTGGTCGGCAATCCTCAGCACAG[cg]ACGCCACTTCCCAAAGTCAACACCAT
CTTTCTTGTACTGAGCACAGCGCTCTGAGAGGCCATCAAGCCCTGCAAGTCACAAAAGA
GAGAAAGGCTTCTTTGTACCTTTGTACCTGATCCATGGGGCTTCTAATAAAGGGAAGGA
GTTCTCCCTTTGCTTAGCTTTCAATCCACTGTGCTTGAGGATTGAAAACAGCCAAGCAT
ATCAGCATTAATCACAACACTGAACCAGAAGACTTAGATTTAATAAATAGTGTTTTGAC
ATACATACTATCTACTCCATATATAGAATAGAAGAAACCAATAGTTAATATGATACTCA
TTTTACAAAGGTGGAAACTGAAGCTCCTAATGGTTAAGCAACTTTACCAAGTTTGAATT
GCTCAAGAGTGACAGAGCTGGGATTCAAATTCTGCTTAGCTAACCCAATGTTGTGAGTT
AATGCTTGTCTACTTGGGCAG

aagaaaatcaaacttaactcggacccagagacattttagtatgtgttggaaactttagc
atctggtcaccatcctccaaagaattatttggattggaactcggtcagagctgtcactc
ttcagctaggaatctaagaggatcatgtcttggatgttacggagtatagacaaccaagt
tccctgccctcaaaagcccgatcacttataagacagcttatggagctttgacagagggc
agcagttgatggcattatcctttgaactcatagcttagttggactcctactggcttgtg
ggaccaaatctttccctaccacagttggctatagcaaaagttgtgaaaaatgccactag
gatatactggtgagggaaaaggaggtccatttgtagttatagtataattgaaaagaaaa
gctctgaagaaaactctagcctactctttttcagcccaaggggaaggcagagcacctgc
tgacagatgctggcgtagcgagccagggcgttggcgttttcctggatagcGAGGCTGGA
TGGACACTGGTCGGCAATCCTCAGCACAG[cg]ACGCCACTTCCCAAAGTCAACACCAT
CTTTCTTGTACTgagcacagcgctctgagaggccatcaagccctgcaagtcacaaaaga
gagaaaggcttctttgtacctttgtacctgatccatggggcttctaataaagggaagga
gttctccctttgcttagctttcaatccactgtgcttgaggattgaaaacagccaagcat
atcagcattaatcacaacactgaaccagaagacttagatttaataaatagtgttttgac
atacatactatctactccatatatagaatagaagaaaccaatagttaatatgatactca
ttttacaaaggtggaaactgaagctcctaatggttaagcaactttaccaagtttgaatt
gctcaagagtgacagagctgggattcaaattctgcttagctaacccaatgttgtgagtt
aatgcttgtctacttgggcag

FIGURE 23iii

47980
GCCTTAGTTTTGGTAATCTGCAAAACCAAGGGCCCTGCCCTGGGTGCTGCTCTCCCAGT
GCAAAGTCCCTAACTTTGGTGTGACCCTTACCAGAGTCAAGGCTGTCTGGGCCTGGCTC
CTTGTGACATCCATCGCCATCTCTCCGAGGGGCTAAGAGGTAGATGCTTTGGGAGGCAG
AGATGCTCCTGCCTGCTGAGGCCTAGCACATGCTGTAGCCTTGGAGCGTAAGGCCCGCC
TGTGGCAGCAACGGCTGCTTGGATGGAGAGCTGCCTGAGGCTGGCAGCCCAGGGCTTCT
GCACTGAAAGGGCTCAGCCTGGCGGCTGCTCAAATACTCTGCCCCCTGCCATGGGGTCA
GAGGCAGGGCAGAAAGGGAGGGTAGGCCATGTGGGTAACAGTTGACAGGGCCACGGGGA
CAGAGCCATGGGGCAGCCGGCCACACTCTGTGAACATGGGGTAGGGATTGCTGCCCAGC
AGGAGGGGGTGTGCAGAGCCAGCCTACCCATCTTCCATTCCTCAGCCTTGTGCGGGCAG
AAAGTCACCAGGCTGCCTTGGCCACAGAA[ct]ACTTACTGAAATGCCCTTGGACAGGG
AGGGGGTCCTAAGGGGGCCTGGCCCGCGCTGGTGCAGGTCTGGACTTGCTCTTGGAGGC
AAGGGGATCCCCAGTGGATTTTCATCTGCAGAGAGGTTCGATTTGCATTTCATACAATC
CAGGGGTCTGTATGGAACTTGGGGAAGGGGTGGTGGAGGAAGGTGGCCAACTGATCAAA
AACAAACAAAAAACAGGGGTATCATTCTTAATTTTGTGACTGCAAAGTCCAGGCCTCAG
GCTTGCTTTGGGTGCCTCCATGGGCATAGACCATGACTTCCAGGCTCTGGCCCAGGCCT
CTCCTTGGGCTCACCTGGGAGTGACATCCACATGCTATGTACTTGCTGGCACCTGCCAA
AGCCTGCTAAAATTAGCTGGAGCTGGCAAGTGGGTCAGGGTATGGAGGGTGCCTTGTCA
GAATGCCAGGTCTCTCGCCAA

gccttagttttggtaatctgcaaaaccaagggccctgccctgggtgctgctctcccagt
gcaaagtccctaactttggtgtgacccttaccagagtcaaggctgtctgggcctggctc
cttgtgacatccatcgccatctctccgaggggctaagaggtagatgctttgggaggcag
agatgctcctgcctgctgaggcctagcacatgctgtagccttggagcgtaaggcccgcc
tgtggcagcaacggctgcttggatggagagctgcctgaggctggcagcccagggcttct
gcactgaaagggctcagcctggcggctgctcaaatactctgccccctgccatggggtca
gaggcagggcagaaagggagggtaggccatgtgggtaacagttgacagggccacgggga
cagagccatggggcagccggccacactctgtgaacatggggtagggattgctgcccagc
aggagggggtgtgcagagccagcctacccatcttccattcctcagccttgtgcgggcAG
AAAGTCACCAGGCTGCCTTGGCCACAGAA[ct]ACTTACTGAAATGCCCTTGGACAGGG
AGGGGgtcctaagggggcctggcccgcgctggtgcaggtctggacttgctcttggaggc
aaggggatccccagtggattttcatctgcagagaggttcgatttgcatttcatacaatc
caggggtctgtatggaacttggggaaggggtggtggaggaaggtggccaactgatcaaa
aacaaacaaaaaacaggggtatcattcttaattttgtgactgcaaagtccaggcctcag
gcttgctttgggtgcctccatgggcatagaccatgacttccaggctctggcccaggcct
ctccttgggctcacctgggagtgacatccacatgctatgtacttgctggcacctgccaa
agcctgctaaaattagctggagctggcaagtgggtcagggtatggagggtgccttgtca
gaatgccaggtctctcgccaa

FIGURE 23jjj

47981
ATAAAAGCCCCAGGCCAGGCCCCGGACACTGGTGTCCTGGGTCACCGTTAGCTCCAGGA
ATAAGTACCCTAGAACCCCTCGAGAGGCTGGACACTGGATAGCCACAGTGAGGAGGGGT
GGTGGGCAGAGGGCCAGTGGCAGGCACAGCTGCCCTAGCCAGGACCCCCAAGGCCCATG
TGCCTCCTTCCAAGGTGCCCCAAGCCTGCTCGCCTTCCCTGCCCCCAGCCTTAGTTTTG
GTAATCTGCAAAACCAAGGGCCCTGCCCTGGGTGCTGCTCTCCCAGTGCAAAGTCCCTA
ACTTTGGTGTGACCCTTACCAGAGTCAAGGCTGTCTGGGCCTGGCTCCTTGTGACATCC
ATCGCCATCTCTCCGAGGGGCTAAGAGGTAGATGCTTTGGGAGGCAGAGATGCTCCTGC
CTGCTGAGGCCTAGCACATGCTGTAGCCTTGGAGCGTAAGGCCCGCCTGTGGCAGCAAC
GGCTGCTTGGATGGAGAGCTGCCTGAGGCTGGCAGCCCAGGGCTTCTGCACTGAAAGGG
CTCAGCCTGGCGGCTGCTCAAATACTCTG[tc]CCCCTGCCATGGGGTCAGAGGCAGGG
CAGAAAGGGAGGGTAGGCCATGTGGGTAACAGTTGACAGGGCCACGGGGACAGAGCCAT
GGGGCAGCCGGCCACACTCTGTGAACATGGGGTAGGGATTGCTGCCCAGCAGGAGGGGG
TGTGCAGAGCCAGCCTACCCATCTTCCATTCCTCAGCCTTGTGCGGGCAGAAAGTCACC
AGGCTGCCTTGGCCACAGAACACTTACTGAAATGCCCTTGGACAGGGAGGGGGTCCTAA
GGGGGCCTGGCCCGCGCTGGTGCAGGTCTGGACTTGCTCTTGGAGGCAAGGGGATCCCC
AGTGGATTTTCATCTGCAGAGAGGTTCGATTTGCATTTCATACAATCCAGGGGTCTGTA
TGGAACTTGGGGAAGGGGTGGTGGAGGAAGGTGGCCAACTGATCAAAAACAAACAAAAA
ACAGGGGTATCATTCTTAATT

ataaaagccccaggccaggccccggacactggtgtcctgggtcaccgttagctccagga
ataagtaccctagaacccctcgagaggctggacactggatagccacagtgaggaggggt
ggtgggcagagggccagtggcaggcacagctgccctagccaggacccccaaggcccatg
tgcctccttccaaggtgccccaagcctgctcgccttccctgcccccagccttagttttg
gtaatctgcaaaaccaagggccctgccctgggtgctgctctcccagtgcaaagtcccta
actttggtgtgacccttaccagagtcaaggctgtctgggcctggctccttgtgacatcc
atcgccatctctccgaggggctaagaggtagatgctttgggaggcagagatgctcctgc
ctgctgaggcctagcacatgctgtagccttggagcgtaaggcccgcctgtggcagcaac
ggctgcttggatggagagctgcctgaggctggcagcccagggcttctgcactgaaaggg
ctcagCCTGGCGGCTGCTCAAATACTCTG[tc]CCCCTGCCATGGGGTCAGAGGCAGgg
cagaaagggagggtaggccatgtgggtaacagttgacagggccacggggacagagccat
ggggcagccggccacactctgtgaacatggggtagggattgctgcccagcaggaggggg
tgtgcagagccagcctacccatcttccattcctcagccttgtgcgggcagaaagtcacc
aggctgccttggccacagaacacttactgaaatgcccttggacagggagggggtcctaa
gggggcctggcccgcgctggtgcaggtctggacttgctcttggaggcaaggggatcccc
agtggattttcatctgcagagaggttcgatttgcatttcatacaatccaggggtctgta
tggaacttggggaaggggtggtggaggaaggtggccaactgatcaaaaacaaacaaaaa
acaggggtatcattcttaatt

FIGURE 23kkk

47986
CAGGCAGGGTCTGCAGTGGTATCACTGTGGGCAGAGCCTGGGGAGGGGGCCAATTCTGT
GCACAGGGCAAGGGCGAGAGGAGGGGCCAGGGATCTAGGGCTCCGGGGAGGGGTCAGCA
GGTCGGGGGGAGGGATCCACGGGGAGGGGTTACCCTGGGTGAAGAAGTGAGCCTTGTAC
TTTCCAGTCCGCACAGCAAAAACCCCACGGACCTCGTCTGGGTAGGACGGGTAGAAGAA
GAGAGACTGCCGAGGGCTCTGGGGGCAGAGTCAGGGGTCACGGGGCGGGGCAGGCCCCA
AGCACTGCACATACCTGGGGCTGCCAGCCCTGGTGGGAGGCCCTGGACGTGCACCGCTT
CTTGCCCACCCAGGAACCTGAGAGGTGGCGCCACTTGGATGCCACTCAGTGCAGGAGGC
ACTGAGGCACAGACTCTCAGGCACTGCCCACACTCACCCCAGGGGAAGGCCAGGACAGG
GGCCAAGGATCTGGGATCAGGGGTCACCGGCCCTACCTTGCCTGTGCCCAGCAGCAGGG
GGCTGAGGTCAAAGCCATCCAAGGTGACA[tc]TGGGCAGTGGGGCCCCAGCCAGGGCT
GCCAGGGTAGGCAGCAGGTCCAGGGAGCTGGCCAGCTCGTGGGTCACGCCTGGGGGCAG
GAGGCTGGTCAGTCACTCAGTTCGCCATCAAGGTTGGGGTGGTGGGGCCAGGGTTCCAA
GGAGAGGGCCTGCGGACTGACCGGGAGCGATATGACCTGGCCAGAAGGCCAAGGCAGGC
TCTCGGACACCGCCCTCGTAGGTCGTTCCCTTTCCACACCGCAAGAGACCGGAGCAGCC
GCCTCGGGACATACGCATGGTCTCAGGTCTGGGACACAGGAGGCGCTCATGAGCCATGG
AGCCACAGCCTCTGAGCCACCGAGGGTGACCAGTGGCCCCACACCTCTAAGTCACAAAG
CTTGCCCGGAGGTGCCCAGCATGAGCCCGGCACCTCCCAGGCCTACCAAGACCAGCTCT
CTGTGCACTGTGTCTCCTGAC

caggcagggtctgcagtggtatcactgtgggcagagcctggggagggggccaattctgt
gcacagggcaagggcgagaggaggggccagggatctagggctccggggaggggtcagca
ggtcggggggagggatccacggggaggggttaccctgggtgaagaagtgagccttgtac
tttccagtccgcacagcaaaaaccccacggacctcgtctgggtaggacgggtagaagaa
gagagactgccgagggctctgggggcagagtcaggggtcacggggcggggcaggcccca
agcactgcacatacctggggctgccagccctggtgggaggccctggacgtgcaccgctt
cttgcccacccaggaacctgagaggtggcgccacttggatgccactcagtgcaggaggc
actgaggcacagactctcaggcactgcccacactcaccccaggggaaggccaggacagg
ggccaaggatctgggatcaggggtcaccggccctaccttgcctgtgcccagcagcaggg
ggctgaggtcAAAGCCATCCAAGGTGACA[tc]TGGGCAGTGGGGCCCCAGCcagggct
gccagggtaggcagcaggtccagggagctggccagctcgtgggtcacgcctgggggcag
gaggctggtcagtcactcagttcgccatcaaggttggggtggtggggccagggttccaa
ggagagggcctgcggactgaccgggagcgatatgacctggccagaaggccaaggcaggc
tctcggacaccgccctcgtaggtcgttccctttccacaccgcaagagaccggagcagcc
gcctcgggacatacgcatggtctcaggtctgggacacaggaggcgctcatgagccatgg
agccacagcctctgagccaccgagggtgaccagtggccccacacctctaagtcacaaag
cttgcccggaggtgcccagcatgagcccggcacctcccaggcctaccaagaccagctct
ctgtgcactgtgtctcctgac

FIGURE 23111

47987
GTCCAGCAATGAGTCACAGACCTATGCACCACCTGCAAAGGAGCCAGAGAAAACAAACG
CCCAGCGCTTTTAGCCTGAAAATGAGAATCTGGTTTGCTGGGGAAGATAAAGGGTGTCG
GAAAATGGCTGTTGGGTAAATCATTGATGTCTGCCACTAGGAATGAAAGGCAAATCAGG
AACTGGCACACATGCTTTCAGGGAGATGGCTGCAAGGGAGAGGGCAAAGACTGGAAGT
TGCTTATGTGGTGCCAGACTATTTGGAAGATCATGGATTGCGGTGTTTGTGTTGTGTGG
TCATCATTTTGTTCTTTGTTTACAGAACAGAGAAAGTGGATTGAACAAGGACGCATTTC
CCCAGTACATCCACAACATGCTGTCCACATCTCGTTCTCGGTTTATCAGAAATACCAAC
GAGAGCGGTGAAGAAGTCACCACCTTTTTTGATTATGATTACGGTGCTCCCTGTCATAA
ATTTGACGTGAAGCAAATTGGGGCCCAACTCCTGCCTCCGCTCTACTCGCTGGTGTTCA
TCTTTGGTTTTGTGGGCAACATGCTGGTC[ag]TCCTCATCTTAATAAACTGCAAAAAG
CTGAAGTGCTTGACTGACATTTACCTGCTCAACCTGGCCATCTCTGATCTGCTTTTTCT
TATTACTCTCCCATTGTGGGCTCACTCTGCTGCAAATGAGTGGGTCTTTGGGAATGCAA
TGTGCAAATTATTCACAGGGCTGTATCACATCGGTTATTTTGGCGGAATCTTCTTCATC
ATCCTCCTGACAATCGATAGATACCTGGCTATTGTCCATGCTGTGTTTGCTTTAAAAGC
CAGGACGGTCACCTTTGGGGTGGTGACAAGTGTGATCACCTGGTTGGTGGCTGTGTTTG
CTTCTGTCCCAGGAATCATCTTTACTAAATGCCAGAAAGAAGATTCTGTTTATGTCTGT
GGCCCTTATTTTCCACGAGGATGGAATAATTTCCACACAATAATGAGGAACATTTTGGG
GCTGGTCCTGCCGCTGCTCAT

gtccagcaatgagtcacagacctatgcaccacctgcaaaggagccagagaaaacaaacg
cccagcgcttttagcctgaaaatgagaatctggtttgctggggaagataaagggtgtcg
gaaaatggctgttgggtaaatcattgatgtctgccactaggaatgaaaggcaaatcagg
aactggcacacatgctttcagggagatggctgcaagggagagggcaaagactggaagt
tgcttatgtggtgccagactatttggaagatcatggattgcggtgtttgtgttgtgtgg
tcatcattttgttctttgtttacagaacagagaaagtggattgaacaaggacgcatttc
cccagtacatccacaacatgctgtccacatctcgttctcggtttatcagaaataccaac
gagagcggtgaagaagtcaccaccttttttgattatgattacggtgctccctgtcataa
atttgacgtgaagcaaattggggcccaactcctgcctccgctctactcGCTGGTGTTCA
TCTTTGGTTTTGTGGGCAACATGCTGGTC[ag]TCCTCATCTTAATAAACTGCAAAAAG
CTGAAGTGCTTGACtgacatttacctgctcaacctggccatctctgatctgctttttct
tattactctcccattgtgggctcactctgctgcaaatgagtgggtctttgggaatgcaa
tgtgcaaattattcacagggctgtatcacatcggttattttggcggaatcttcttcatc
atcctcctgacaatcgatagatacctggctattgtccatgctgtgtttgctttaaaagc
caggacggtcacctttggggtggtgacaagtgtgatcacctggttggtggctgtgtttg
cttctgtcccaggaatcatctttactaaatgccagaaagaagattctgtttatgtctgt
ggcccttattttccacgaggatggaataatttccacacaataatgaggaacattttggg
gctggtcctgccgctgctcat

FIGURE 23mmm

47994
CGGGGTCCCAGAAGGTGGTTTAAGGACTGGTGTGGACACACACAGTTCTGTTGTCTGCC
CAGCGGAGGGGCCTCACGGGGGGCCGTTGGGAGCCAGATTGTCAGCTTTTGGATTTACC
AGCTGTGGGTGGCAGTGGGCGTGTAACTCAGCATCTTGCTGCCTCAGTTTCTCTCATCT
GTAAAGTGGGGATAATAACATTTACCTCATAAAGTTCCTGCGAGGATTCGATGACTTGA
TACATCAGTTGCTTAGCACAGGGCTCAGCACTCAGTACATGTTCCCTGTCAGGAAGGCA
GGGAGGCCTCACTGGCAGCATCAGGACATGGACATCAGGACATACACCGTGGCTCTCA
GGGAAAGGAAAAAGACCCTCTCCCAGGTGTACAAGCTCGATTCTAAACCTCATGGGACC
CTGCATTGTTCGCTCCCTCATTCATTCACTAGTCCATGCGTGTACTCAGTAGTGGCATA
AGCAGACTGCTCGGGTCGGACCTGAATTAGCCTCACCCACTCTCCTCCTACACTGTCCC
TCCCCAGGGCACATTCGCCTCCCAGGTGA[ct]GCTGGAGGGGGACAAGTTGAAAGTGG
AGCGGGAGATCGATGGGGGCCTGGAGACCCTGCGCCTGAAGCTGCCAGCTGTGGTGACA
GCTGACCTGAGGCTCAACGAGCCCCGCTACGCCACGCTGCCCAACATCATGGTGAGCCC
CTGGCCAGCGGGCACTGAGGGCCTGGGGGTGGCAAGCACATTGCCAGCCCAGTGCCCCC
CGGTGGTCGCACGTGGGGAGGGAAGGATCCAAAGGAGGTCTCGTGCACAGGAAGCCGTC
ACCTGGAGTTTGGCTGATAGAGAGAGTTTGCTGGGTCATCTCTGCCAATACTGAGAGTT
CATGGGGGCTGCTTTGGCTAGCAGGGAGGGCTTGCTGGTATCTAGGCCAGTAGAAAGCC
TTCGCTGGGCAGCAGAAGGTGTTCCCTTTGTCATTCCAGCCAGTGGAACAAGTTCACTG
GGTCATCTAGGTTCATTAGGG

cggggtcccagaaggtggtttaaggactggtgtggacacacacagttctgttgtctgcc
cagcggaggggcctcacggggggccgttgggagccagattgtcagcttttggatttacc
agctgtgggtggcagtgggcgtgtaactcagcatcttgctgcctcagtttctctcatct
gtaaagtggggataataacatttacctcataaagttcctgcgaggattcgatgacttga
tacatcagttgcttagcacagggctcagcactcagtacatgttccctgtcaggaaggca
gggaggcctcactggcagcatcaggacatggacatcaggacatacaccgtggctctca
gggaaaggaaaaagaccctctcccaggtgtacaagctcgattctaaacctcatgggacc
ctgcattgttcgctccctcattcattcactagtccatgcgtgtactcagtagtggcata
agcagactgctcgggtcggacctgaattagcctcacccactctcctcctacactgtccc
TCCCCAGGGCACATTCGCCTCCCAGGTGA[ct]GCTGGAGGGGGACAAGTTGAAAGTGG
AGCgggagatcgatgggggcctggagaccctgcgcctgaagctgccagctgtggtgaca
gctgacctgaggctcaacgagccccgctacgccacgctgcccaacatcatggtgagccc
ctggccagcgggcactgagggcctgggggtggcaagcacattgccagcccagtgccccc
cggtggtcgcacgtggggagggaaggatccaaaggaggtctcgtgcacaggaagccgtc
acctggagtttggctgatagagagagtttgctgggtcatctctgccaatactgagagtt
catgggggctgctttggctagcagggagggcttgctggtatctaggccagtagaaagcc
ttcgctgggcagcagaaggtgttccctttgtcattccagccagtggaacaagttcactg
ggtcatctaggttcattaggg

FIGURE 23nnn

47997

TACTAAAAATATATAAATTAGCTGGGTGTGGTGGCATGTGCCTGTAATCCCAGGTACTT
GGGAGGCCAAGGCAGGAGTATTGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAG
ATCGTGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCATCTCAAAATAAATAA
ATAAATATAATAAAAATAAATAAACAAATAAGCTTCCTTTTGCTCATTGACCCCAGAAT
CCCAGAGAAACCACACGTCCCAGCAACCCTCGTGGCAGAATAAGCCACAGAAAACAGCC
CACCCTAAGTGCCTCGCCTCCAGCAACTGAAGTTGCACGAGTCAGCACGTGCCCTTCTG
TGGACCTCAGAATAGATCCCTTCATACAAGGGCTGCAGGAGAAAGCAGGACTCCCAGCA
ATCTCTGGGGTCTGAGCTGGCCTGGCAAGCTGCCTCTGGGGCTGCCAGGAACTGCTATC
TCTCTGCACAGAGGTCCAATCCATACCTGCGTTGCAAAGATGGCTCTCTTCATCATAGT
GAAGTCTTCCTTATCCAGCATCTTGTTCA[ct]GTCGGGAAGGCTCCCACTGCAAGGCA
AGCAGGGGGCATGCATGTGAGAACGGAGTAATGAGAGGGGTTAGTCAGGGCCTAGGAGG
GCACAGGGCTGAGGGTGGGGCACTCACACCAGTAAGGATTCATAAAGCTTCCTCCCGAA
CTTTTCCTTCACCGTGTTGGCCGTGTCCCTGGAGGAAGCAGAGCAACAGGGTCACATAC
ACACCAGCTGCCATTTACTGTTAGGCTTCTTTAGTTAGTTTGTTTGTTTATTTTGAGAC
GGAGTTTGGCTCTTGTTGCCCAGGCTGGAATGCAATGGCGTGATCTCGGCTCACTGCAA
CCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTA
CAGGCATGAGCCACCGCGCCCGGCTAATTTTCTATTTTTAGTAGAGACGGGGTTTCTCC
ATGTTGGTCAGGCTGGTCTCA

tactaaaaatatataaattagctgggtgtggtggcatgtgcctgtaatcccaggtactt
gggaggccaaggcaggagtattgcttgaacccaggaggcagaggttgcagtgagccgag
atcgtgccactgcactccagcctgggcgacagagcgagactccatctcaaaataaataa
ataaatataataaaaataaataaacaaataagcttccttttgctcattgaccccagaat
cccagagaaaccacacgtcccagcaaccctcgtggcagaataagccacagaaaacagcc
caccctaagtgcctcgcctccagcaactgaagttgcacgagtcagcacgtgcccttctg
tggacctcagaatagatcccttcatacaagggctgcaggagaaagcaggactcccagca
atctctggggtctgagctggcctggcaagctgcctctggggctgccaggaactgctatc
tctctgcacagaggtccaatccatacctgcgttgcaaagatggctctcttcatcatagt
gaagTCTTCCTTATCCAGCATCTTGTTCA[ct]GTCGGGAAGGCTCCCACTGCAAGGCa
agcagggggcatgcatgtgagaacggagtaatgagaggggttagtcagggcctaggagg
gcacagggctgagggtggggcactcacaccagtaaggattcataaagcttcctcccgaa
cttttccttcaccgtgttggccgtgtccctggaggaagcagagcaacagggtcacatac
acaccagctgccatttactgttaggcttctttagttagtttgtttgtttattttgagac
ggagtttggctcttgttgcccaggctggaatgcaatggcgtgatctcggctcactgcaa
cctctgcctcccaggttcaagcaattctcctgcctcagcctcccgagtagctgggatta
caggcatgagccaccgcgcccggctaattttctatttttagtagagacggggtttctcc
atgttggtcaggctggtctca

FIGURE 23ooo

48083
CCCTCCCCACAGTACTGTGCAGCCCTGGAATCCCTGATCAACGTGTCAGGCTGCAGTGC
CATCGAGAAGACCCAGAGGATGCTGAGCGGATTCTGCCCGCACAAGGTCTCAGCTGGGG
TAAGGCATCCCCCACCCTCTCACACCCACCCTGCACCCCCTCCTGCCAACCCTGGGCTC
GCTGAAGGGAAGCTGGCTGAATATCCATGGTGTGTGTCCACCCAGGGGTGGGGCCATTG
TGGCAGCAGGGACGTGGCCTTCGGGATTTACAGGATCTGGGCTCAAGGGCTCCTAACTC
CTACCTGGGCCTCAATTTCCACATCTGTACAGTAGAGGTACTAACAGTACCCACCTCAT
GGGGACTTCCGTGAGGACTGAATGAGACAGTCCCTGGAAAGCCCCTGGTTTGTGCGAGT
CGTCCCGGCCTCTGGCGTTCTACTCACGTGCTGACCTCTTTGTCCTGCAGCAGTTTTCC
AGCTTGCATGTCCGAGACACCAAAATCGAGGTGGCCCAGTTTGTAAAGGACCTGCTCTT
ACATTTAAAGAAACTTTTTCGCGAGGGAC [ag] GTTCAACTGAAACTTCGAAAGCATCA
TTATTTGCAGAGACAGGACCTGACTATTGAAGTTGCAGATTCATTTTTCTTTCTGATGT
CAAAAATGTCTTGGGTAGGCGGGAAGGAGGGTTAGGGAGGGGTAAAATTCCTTAGCTTA
GACCTCAGCCTGTGCTGCCCGTCTTCAGCCTAGCCGACCTCAGCCTTCCCCTTGCCCAG
GGCTCAGCCTGGTGGGCCTCCTCTGTCCAGGGCCCTGAGCTCGGTGGACCCAGGGATGA
CATGTCCCTACACCCCTCCCCTGCCCTAGAGCACACTGTAGCATTACAGTGGGTGCCCC
CCTTGCCAGACATGTGGTGGGACAGGGACCCACTTCACACACAGGCAACTGAGGCAGAC
AGCAGCTCAGGCACACTTCTTCTTGGTCTTATTTATTATTGTGTGTTATTTAAATGAGT
GTGTTTGTCACCGTTGGGGAT

ccctccccacagtactgtgcagccctggaatccctgatcaacgtgtcaggctgcagtgc
catcgagaagacccagaggatgctgagcggattctgcccgcacaaggtctcagctgggg
taaggcatcccccaccctctcacacccaccctgcaccccctcctgccaaccctgggctc
gctgaagggaagctggctgaatatccatggtgtgtgtccacccaggggtggggccattg
tggcagcagggacgtggccttcgggatttacaggatctgggctcaagggctcctaactc
ctacctgggcctcaatttccacatctgtacagtagaggtactaacagtacccacctcat
ggggacttccgtgaggactgaatgagacagtccctggaaagcccctggtttgtgcgagt
cgtcccggcctctggcgttctactcacgtgctgacctctttgtcctgcagcagttttcc
agcttgcatgtccgagacaccaaaatcgaggtggcccagtttGTAAAGGACCTGCTCTT
ACATTTAAAGAAACTTTTTCGCGAGGGAC [ag] GTTCAACTGAAACTTCGAAAGCATCA
TTATTTGCAGAGACAGGACCtgactattgaagttgcagattcatttttctttctgatgt
caaaaatgtcttgggtaggcgggaaggagggttagggaggggtaaaattccttagctta
gacctcagcctgtgctgcccgtcttcagcctagccgacctcagccttccccttgcccag
ggctcagcctggtgggcctcctctgtccagggccctgagctcggtggacccagggatga
catgtccctacacccctcccctgccctagagcacactgtagcattacagtgggtgcccc
ccttgccagacatgtggtgggacagggacccacttcacacacaggcaactgaggcagac
agcagctcaggcacacttcttcttggtcttatttattattgtgtgttatttaaatgagt
gtgtttgtcaccgttggggat

FIGURE 23ppp

48088
GTACATACACACCCATGTGATACATATACACATACCCATAGTATACAGGTAACATAAAA
TTATACACACACAACACAAACACATATTATGCACATACGCACATAACACACACACACAC
ACCCACATACAGGCATTGTGAACTAGACACATCACCTTACAATCTGTGGTTTACTGGAA
GGACATGGAACAAAACCCCCCCAGCCACAGCGTGGAAGTGCCCTCTCCAGGCACAAGAT
TCTGCCTCCATGGGGCGTGGTAGCAGCATTGCCCACCCACCCAGGGCTGAGTGAGCAGG
CCTGCCCCACACTGCGCCCATGCACAGCCACTCCAGGCTGCCTCCCACACTGCCTGCAA
GGACCCCAGTGGGGACTGCAAACGGGAAGTCTGCATCCAGGGCCCCAGGGAGGGCAGGT
GGGGCTCTGGAGTATAGCACTTTCTAGAAGGGAAGCACCCTCTTGGTTCTGAACGTAAG
TGGGTCTGCTCACAGGGAGGGGCGTGCAGCCACCCCAGGACCCCAGCTGTCCAAGGAGC
CAGGGAAAACGCACCCACGGGGCACCTAC[ct]GCTGGGAGCGCAAAGAAGGAGATGGC
AAAGACAGAGAAGCAGGAGGCGATGGTCTTCCCGACCCACGTCTGGGGCACCTTGTCCC
CATAGCCGATGGTGGTGACTGTGACCTGCAGGGAGAGGGACAGTGGTCAGCCACGGATG
GGACTGGAGCCTCGGGAGGGCCAACTGCCTAACCCAAACCCACCACTCTGATGAGCGGA
GAGGCCGGCAAGAGACCCTGACCACCAGGACGACCCCGTGTGACTCGGCGAAAGCACCA
GGAACAGAGCCGCGGGATGGCACATGTCTCCCAGGCTCTCGGCGTCACACACAAGGTAT
GTCCCACCAGCACATGTAAGGAGCCCAGCACCCACGAAGGGCCAGGCCTGCTGGCTGGG
AACGTGGGCCTGGGAGCTCGCCCCACACCGGCTGCCTCATCTGCCTGCCTGTCCCCAGG
AGGCTGGGCCCCTGGGCCACC

gtacatacacacccatgtgatacatatacacatacccatagtatacaggtaacataaaa
ttatacacacacaacacaaacacatattatgcacatacgcacataacacacacacacac
acccacatacaggcattgtgaactagacacatcaccttacaatctgtggtttactggaa
ggacatggaacaaaacccccccagccacagcgtggaagtgccctctccaggcacaagat
tctgcctccatggggcgtggtagcagcattgcccacccacccagggctgagtgagcagg
cctgccccacactgcgcccatgcacagccactccaggctgcctcccacactgcctgcaa
ggaccccagtggggactgcaaacgggaagtctgcatccagggccccagggagggcaggt
ggggctctggagtatagcactttctagaagggaagcaccctcttggttctgaacgtaag
tgggtctgctcacagggaggggcgtgcagccaccccaggaccccagctgtccaaggagC
CAGGGAAAACGCACCCACGGGGCACCTAC[ct]GCTGGGAGCGCAAAGAAGGAGATGGC
AAAGacagagaagcaggaggcgatggtcttcccgacccacgtctggggcaccttgtccc
catagccgatggtggtgactgtgacctgcagggagagggacagtggtcagccacggatg
ggactggagcctcgggagggccaactgcctaacccaaacccaccactctgatgagcgga
gaggccggcaagagaccctgaccaccaggacgaccccgtgtgactcggcgaaagcacca
ggaacagagccgcgggatggcacatgtctcccaggctctcggcgtcacacacaaggtat
gtcccaccagcacatgtaaggagcccagcacccacgaagggccaggcctgctggctggg
aacgtgggcctgggagctcgccccacaccggctgcctcatctgcctgcctgtccccagg
aggctgggcccctgggccacc

FIGURE 23qqq

48094
AGCCTGGGTGACAAGAGCAAAACTCCACATCAAAAAAAATAATAATAAATAAATTAATT
AATTAATTAAATAAAACAAGAGCTTTTCTTTTTGCTTAATAAGAGAGAGTGGTGGTGGT
GCTTTTTTATTCCTGAAGATGGGAAGTCCTCTTTTGCCCACTAACCTCAGAAGAAAGGG
ATGAGGTGTACCGTACAGGGGCAGTCACCTTCTCCTCTGTTTAGCTTCCATTTTGGCCT
CATGTCTACCCCAAAGTTGTAGCTTAGATGGGGGGAAAATTCAGAATTTTGCATAGACC
ATAGGTAGCACCCCCTAGAAAAAGAATGTTTCTCCCCAGATGTCTCCCACTAGTACCCT
AACCATCTGCTTGTCTGTCTAGTGAGGACCCTTGGAGGGCTGCTAAAATGATCAAGGGT
TACATGCAGCAACACAACATCCCCCAGAGGGAGGTGGTCGATGTCACCGGCCTGAACCA
GTCGCACCTCTCCCAGCATCTCAACAAGGGCACCCCTATGAAGACCCAGAAGCGTGCCG
CTCTGTACACCTGGTACGTCAGAAAGCAA[ct]GAGAGATCCTCCGACGTAAGTGTTTT
CATCCTGCCTCTGCCTCAACCTGAAGTGACCTTTGCCCTCTCACCCCATTGGCTGCCTC
AGTTTCCCTTTCATCGACAAGGCCTTGTGAGCACTTGGCAGATATGAGGAAGGTGGCAA
GTAGATTTGGCCTTGGTGGTTGCTGTACAATGGATTGGCTTCTGTCATGTTCTTCAGTC
ACAGCCCCCTTGCTACCCAGCCAGTTGCTCTGAGGAGCCTGTCAGTGTATGCAGCATAC
CTTAAACTTTTTGGCCCCTCCTTCCACCTCCTTCTCTTTGAAACCAAGTAGGTGACAGA
GTGAAATGTCTTCCCTGAGAGAAAACCCAGCATCTCCCCTTGATACGTGACCATCAGTC
AATTTCCAAAGAAGACATTTCGTTGCAGTCAATAATATTGATTACTATTACTGTTAATT
TCCTCCTCTCTGGAAAAAGTA

agcctgggtgacaagagcaaaactccacatcaaaaaaaataataataaataaattaatt
aattaattaaataaaacaagagcttttctttttgcttaataagagagagtggtggtggt
gcttttttattcctgaagatgggaagtcctcttttgcccactaacctcagaagaaaggg
atgaggtgtaccgtacaggggcagtcaccttctcctctgtttagcttccattttggcct
catgtctaccccaaagttgtagcttagatggggggaaaattcagaattttgcatagacc
ataggtagcaccccctagaaaaagaatgtttctccccagatgtctcccactagtaccct
aaccatctgcttgtctgtctagtgaggacccttggagggctgctaaaatgatcaagggt
tacatgcagcaacacaacatcccccagagggaggtggtcgatgtcaccggcctgaacca
gtcgcacctctcccagcatctcaacaagggcacccctatgaagacccagaaGCGTGCCG
CTCTGTACACCTGGTACGTCAGAAAGCAA[ct]GAGAGATCCTCCGACGTAAGTGTTTT
CATCCTGCCTCtgcctcaacctgaagtgacctttgccctctcaccccattggctgcctc
agtttccctttcatcgacaaggccttgtgagcacttggcagatatgaggaaggtggcaa
gtagatttggccttggtggttgctgtacaatggattggcttctgtcatgttcttcagtc
acagcccccttgctacccagccagttgctctgaggagcctgtcagtgtatgcagcatac
cttaaactttttggcccctccttccacctccttctctttgaaaccaagtaggtgacaga
gtgaaatgtcttccctgagagaaaacccagcatctccccttgatacgtgaccatcagtc
aatttccaaagaagacatttcgttgcagtcaataatattgattactattactgttaatt
tcctcctctctggaaaaagta

FIGURE 23rrr

48095
CTGACTTAGCTGGGTGATATTGGGCGGGTTTCCTCTCTCTGGCCGTTTCCCACACCTGC
AGGCTGGGAGTGGTGCCTGCTGCCTCCTGACAGTGCTGCAGTGAGCATCAAGTGAGACA
AGCCCATGAAAACCCTCTGCAGCCCCAGAATGCCACGGAAATGCAGCATTATTGTATTG
AGCTTTGCTTTGAGTTTATTATATCATCAAACATATTATTAAATGACTGAGTTGGGTGG
GGGGTTGGTCAAGAGGGCCTATACAAGACCCCAGGATTCTGTGGGACCTGAGATTCTAG
AATTCTGCCACCCTGATTCCAAAGCAAGAGAAGAGTCTCTGACATGATCAGGGCCAGAA
AACTGGCTGGAGAGGCAGACAGTACAGTGCGTTCATATAAATGACTCTAATTCAGGTGG
TGGCGTGAGACTGTGGGCATGTGTGATGTGCAACAGAGCAGGCTGGTGTCCATAAGCCA
ACGATGGCACAGTACTCACCTTCTGGGGGGCATTGATGACTCCAGTGTTGTAGCCAAAC
TGCAGGGAGCCAAGCACTGCTCCTCCCAC[ga]GCCAGCATGAGGCGACCCGTCAGCTT
CTGCGGAGAAACAAACCACACTGTTATAGGCGTGTCTGGGAGCAGGTTACTACAGGGCA
GGGCCTGGACTGGCAAGTTTCTGTGTTCAGATATCTTGCCTGACTCTTGGCACCACACC
AGTCTTTCTCCCAGGAAACTTGGCCAATTCCTGACCTTAGGTGCCCAAACCAGCCTAGC
TGACTTCAAGATACTGGGCTGGCCGGGCCATTTCCTGGGGAGAGAGGGGAAGTATGATC
TTCTCTCTCTGTAGCCAGGTCTCAGAGAGGGAGAGGCTTTGGATTCTTGGGGGTCTCAT
TTCCCTGGTGGAGCCATGCCTAGGGTCTGGTGGTTCTAGACTCTCTGACTGGGAGGCCC
AGGAACCAGCCCTCCTATGCGAGGGGGCCCAAATTACTTGGTAGGAATAGCACAGATAT
AGATAGGAGAAGCACCCTGGA

ctgacttagctgggtgatattgggcgggtttcctctctctggccgtttcccacacctgc
aggctgggagtggtgcctgctgcctcctgacagtgctgcagtgagcatcaagtgagaca
agcccatgaaaaccctctgcagccccagaatgccacggaaatgcagcattattgtattg
agctttgctttgagtttattatatcatcaaacatattattaaatgactgagttgggtgg
ggggttggtcaagagggcctatacaagaccccaggattctgtgggacctgagattctag
aattctgccaccctgattccaaagcaagagaagagtctctgacatgatcagggccagaa
aactggctggagaggcagacagtacagtgcgttcatataaatgactctaattcaggtgg
tggcgtgagactgtgggcatgtgtgatgtgcaacagagcaggctggtgtccataagcca
acgatggcacagtactcaccttctggggggcattgatgactccagtgttgtagccaaac
tgcagGGAGCCAAGCACTGCTCCTCCCAC[ga]GCCAGCATGAGGCGACCCGTCAGCtt
ctgcggagaaacaaaccacactgttataggcgtgtctgggagcaggttactacagggca
gggcctggactggcaagtttctgtgttcagatatcttgcctgactcttggcaccacacc
agtctttctcccaggaaacttggccaattcctgaccttaggtgcccaaaccagcctagc
tgacttcaagatactgggctggccgggccatttcctggggagagaggggaagtatgatc
ttctctctctgtagccaggtctcagagagggagaggctttggattcttgggggtctcat
ttccctggtggagccatgcctagggtctggtggttctagactctctgactgggaggccc
aggaaccagccctcctatgcgagggggcccaaattacttggtaggaatagcacagatat
agataggagaagcaccctgga

FIGURE 23sss

48097
CATAATTTTTCTCAAACTCGGCGGACGGTTCGTGTTTGAAAGAGAAGTTGCCATTGATG
CTGAGCGGCGGGCTGAGGGGTCCATCAAAGGAAGGGCTGGTGCAATCAGTCAGAGGGCT
TTCAAAGAAGGGCTCCAGCGCTGCGCTGTAGGCGTGCGGCGGAGGCTTAACGTGGAAGA
CATGGGAGCTGTCCATGGTACCGTAAGGCGGACTGGGCAGCCCAGGCGACTGGTAGGAG
TAGGGGTGTACAGGGAAGGAAGCGCTGGCCGTCGGCAGGTGGGGGGGCATGTCCTGGTT
CTGCTCAGGCAGAAAAGTCCGAGGATTGAGTTGCAGGCAGCCCGCAACCAGGTTGGTGG
TGGGTTGGGATAAGCCCTTGCAAAGCGTCTGAACGAAGGAGACCAGGTCTGGGCTTTTG
CCTGAGCGCAGGATCTCCGACAGAGCCCAGATGTAGTTCTTGGCCAAGCGCAGAGTCTC
GATTTTGGACAGCTTCTGCGTCTTAGAATAGCAAGGCACCACCTTGCGCAGGTTGTCTA
GCGCCGCGTTCAGTCCGTGCATGCGGTTC[ca]GCTCCCGGGCGTTAGCCTTCATGCGT
CTCAATTTAAAACGCTCCAGGCGAGCCTTAGTCATCTTCTTCTTTTTGGGGCCGCGTCT
CTTGGGCTTTTGATCGTCATCCTCCTCTTCCTCTTCTTCCTCCTCTTCCAGGTCCTCAT
CTTCGTCCTCCTCCTCTCCCCCGTTCCTCAGTGAGTCCTCCTCTGCGTTCATGGTTTCG
AGGTCGTCCTCCTTCTTGTCTGCCTCGTGCTCCTCGTCCTGAGAACTGAGACACTCGTC
TGTCCAGCTTGGAGGACCTTGGGGCTGAGGCTCGCCCATCAGCCCACTCTCGCTGTACG
ATTTGGTCATGTTTCGATTTCCTACATTCAACAAGGGAGAGGCAAACAGAAAGAAAAGC
AGAAAAACGCTATATTCAAAAGCCAGATACGCCTTCAGCTTCCACTCCCTAAACCTGTA
CAAATGCTTGCGAAAAGTACC

cataatttttctcaaactcggcggacggttcgtgtttgaaagagaagttgccattgatg
ctgagcggcgggctgaggggtccatcaaaggaagggctggtgcaatcagtcagagggct
ttcaaagaagggctccagcgctgcgctgtaggcgtgcggcggaggcttaacgtggaaga
catgggagctgtccatggtaccgtaaggcggactgggcagcccaggcgactggtaggag
taggggtgtacagggaaggaagcgctggccgtcggcaggtgggggggcatgtcctggtt
ctgctcaggcagaaaagtccgaggattgagttgcaggcagcccgcaaccaggttggtgg
tgggttgggataagcccttgcaaagcgtctgaacgaaggagaccaggtctgggcttttg
cctgagcgcaggatctccgacagagcccagatgtagttcttggccaagcgcagagtctc
gattttggacagcttctgcgtcttagaatagcaaggcaccaccttgcgcaggttgtcta
gcgcCGCGTTCAGTCCGTGCATGCGGTTC[ca]GCTCCCGGGCGTTAGCCTTCATGCGt
ctcaatttaaaacgctccaggcgagccttagtcatcttcttctttttggggccgcgtct
cttgggcttttgatcgtcatcctcctcttcctcttcttcctcctcttccaggtcctcat
cttcgtcctcctcctctcccccgttcctcagtgagtcctcctctgcgttcatggtttcg
aggtcgtcctccttcttgtctgcctcgtgctcctcgtcctgagaactgagacactcgtc
tgtccagcttggaggaccttggggctgaggctcgcccatcagcccactctcgctgtacg
atttggtcatgtttcgatttcctacattcaacaagggagaggcaaacagaaagaaaagc
agaaaaacgctatattcaaaagccagatacgccttcagcttccactccctaaacctgta
caaatgcttgcgaaaagtacc

FIGURE 23ttt

53412
GGATTTATCTAGTATAACAAACCATCGGTCTGATAATACATATCTGATAGTGTTGCTGT
GAATATAATTGAGGTAATACATGTAAAAGAGCTGGCACACAAAAAGAAGCTCAAAAAAT
TGTTCTTTCCTTACCAGGTGTTGCCCTGGTTCCTGCCATATCGCTCCCCAAAGGTGCTG
TAGGAGCCATCATAGTGTTTGTAGTTCAACTGTCTCTGGTAACCTGGAAAGGAAGATTA
ACGAAACAGCACAATGGATTAATGTGCATGCTGAGGGTGGAGAAATTACTAAAAGTACC
TTGGCTTCTCTTGTGACATTTCTTAAATTTTGTTGTCATAGATTAGGAGTTTCTGAGCC
TTAAATATTTTATTGGAGGTTGGAGAGTGGATAGTTTCCTTGAAATTAACTATCATAGC
AGCTATCATAGTGAGCTAAGCTAATGTATCATAATATTCATAAGTAACTGAAACCTACT
GGGAAATCCAGTTGAAATAACATTCAAGTTTTCCCTTACTCAAGTAATCACTCACCAGT
GTTGAGATAGCCAATGGCCTTGGACTTGA[ct]CTCTGGAGTAAGCTGCTGTGTTTCAT
TTAGATAATCCAGTACATAGATGTTAGGAGCAAAGAGGACCATATTCTGCTCTCCACAG
CCATAGGGCATCTGGAGAAGATTTTGTGTGTTTTGCATGGCAGAGCCTAATATGTCTCC
TAGAGAATGGGAGAGATGGGAAGTCATAAAGCTTGGAGATTATCATCTATCAAAGTCAT
TAAGCAGAAATAATTAGTTGAGCTTAGAAATTGAGAATTTTTAGGAAGGATGATTCTTC
CAGGGATAGAAGTATGATTGAAAGCAATAAACAAGCCCAAAGAAGAAGAGAAGAAAGAA
GTTAAAATTATAGTATTATTTTTAGTAAATATTTATGGGAAATAAAAATAGTATAATAG
AAGCTGTTAATGCCCGGATCCACTAGGGGCTGGAGACTCACCCAAAACTGAGACAGAAG
CTCGGGCAGATTCTTCTACCA

ggatttatctagtataacaaaccatcggtctgataatacatatctgatagtgttgctgt
gaatataattgaggtaatacatgtaaaagagctggcacacaaaaagaagctcaaaaaat
tgttctttccttaccaggtgttgccctggttcctgccatatcgctccccaaaggtgctg
taggagccatcatagtgtttgtagttcaactgtctctggtaacctggaaaggaagatta
acgaaacagcacaatggattaatgtgcatgctgagggtggagaaattactaaaagtacc
ttggcttctcttgtgacatttcttaaattttgttgtcatagattaggagtttctgagcc
ttaaatattttattggaggttggagagtggatagtttccttgaaattaactatcatagc
agctatcatagtgagctaagctaatgtatcataatattcataagtaactgaaacctact
gggaaatccagttgaaataacattcaagttttcccttactcaagtaatCACTCACCAGT
GTTGAGATAGCCAATGGCCTTGGACTTGA[ct]CTCTGGAGTAAGCTGCTGTGTTTCAT
TTAGATAATCCAGTacatagatgttaggagcaaagaggaccatattctgctctccacag
ccatagggcatctggagaagattttgtgtgttttgcatggcagagcctaatatgtctcc
tagagaatgggagagatgggaagtcataaagcttggagattatcatctatcaaagtcat
taagcagaaataattagttgagcttagaaattgagaatttttaggaaggatgattcttc
cagggatagaagtatgattgaaagcaataaacaagcccaaagaagaagagaagaaagaa
gttaaaattatagtattatttttagtaaatatttatgggaaataaaaatagtataatag
aagctgttaatgcccggatccactaggggctggagactcacccaaaactgagacagaag
ctcgggcagattcttctacca

FIGURE 23uuu

53415
GAAAACAGTCTGGGTTCCCTGTGGAACATCATTTCTCAAAACTGTATTTTGGGGCTTGC
TTTCTCATTTTTCCTTTCCATTTCAGATATCCTTACTGCTGTCTTTGGGCTCTTTAAAC
ACTGCCTTTTTTCCTTTTTCGATCACACCCAAAAACTTTTCTCAAAAATTACATGTAAA
TTTAAAAATTTACAAATTAAATTTAAAATTGAAATTTTAAAAATCCCGACTCTCCCTAA
TTTCAGGAAGCATGCATTTATTATACATAACAAGACGTGAAAGCCGCAAGAGTTTCAGC
CTAAACACTGAAGACCCCGCGAAGTGAATCCAGCTGCTGCTCTACAAGCAGCAACAACA
ACTGGGAAGCCTTCTCAGCTACACTTCGGGGCACTGGTCCAACCCCACGCAAAATCCCT
CGTTTCCCTTAGCGTGGTAAGACGGAGCCTGACCTGAGCTCCAACTGTCCTATCTTTTT
CAAATGTTTCAAACTTACTGCCTTTGTTCAGCAGAACCACGGGCACGGTGATGATGGTG
ACAAGCGCAGCAGCACCCAGCAGTCCCAG[ga]AGAACCTTCCACGGTGTCTGCAAGCC
GAGCAGATCAAGTCCAATTAGAGGGAAGCGTGTGGCCCCAGTTTCCGTAGGAGGGTCGG
GGCTGCTCCAGAGGCAGCAGGATTTGCAGGTGGGAGTGCGTTAGAAGAGGGAGACCGCG
GGCTGGGGGTGGGGGTGGCGTCTGGAGTGCGCCAGTTGGAGTTCTCTAAGGCGGGTGCC
CTTGAACTTGTGCCTTCAGAGCACATTAGCGTTGGTTTCTCTACCCCTGCCCGGGTTCG
GGCGTGCGTTCTGTGAGTGGCTCTCCGGGACATTCAAAGCTCGACGCCAGGGTCCTAGC
AGAAGCCAGGGTCCGAAAGCTAAGCGAGAGCTCTGGGACGTCCCTTCACCTGTCAGAGG
GTGGCCTTGGGGCTTCCGCCTAAGGGGAGTCCCTGGTCCGGTTTCGCCAGCTTTTGGGC
CATTTGGGGAGTTTGGCGAAG

gaaaacagtctgggttccctgtggaacatcatttctcaaaactgtattttggggcttgc
tttctcatttttcctttccatttcagatatccttactgctgtctttgggctctttaaac
actgccttttttcctttttcgatcacacccaaaaacttttctcaaaaattacatgtaaa
tttaaaaatttacaaattaaatttaaaattgaaattttaaaaatcccgactctccctaa
tttcaggaagcatgcatttattatacataacaagacgtgaaagccgcaagagtttcagc
ctaaacactgaagaccccgcgaagtgaatccagctgctgctctacaagcagcaacaaca
actgggaagccttctcagctacacttcggggcactggtccaaccccacgcaaaatccct
cgtttcccttagcgtggtaagacggagcctgacctgagctccaactgtcctatcttttt
caaatgtttcaaacttactgcctttgttcagcagaaccacgggcacggtgatgatggtg
aCAAGCGCAGCAGCACCCAGCAGTCCCAG[ga]AGAACCTTCCACGGTGTCTGCAAGCC
GAgcagatcaagtccaattagagggaagcgtgtggccccagtttccgtaggagggtcgg
ggctgctccagaggcagcaggatttgcaggtgggagtgcgttagaagagggagaccgcg
ggctgggggtgggggtggcgtctggagtgcgccagttggagttctctaaggcgggtgcc
cttgaacttgtgccttcagagcacattagcgttggtttctctacccctgcccgggttcg
ggcgtgcgttctgtgagtggctctccgggacattcaaagctcgacgccagggtcctagc
agaagccagggtccgaaagctaagcgagagctctgggacgtcccttcacctgtcagagg
gtggccttggggcttccgcctaaggggagtccctggtccggtttcgccagcttttgggc
catttggggagtttggcgaag

FIGURE 23vvv

53418
AGAGGCACAAGAAATTACCTTGAAAAAATGGAATTCAGTGACTGCCATCTAGGAAAGAC
AGTGATACTGTCCAGCAGCATGCAGTTCCGAGAGCTCAACTCTTAGGCCACCCTCCCTC
CACTCTACTCTAGGAACAAGGAGCATTAGGTCTGTTTTCTCTCCATACACCTCAATCGC
TCGTCCTCTCGTCTTATTAAAACACAGACACAGAACCAAACTTTTTGACAGTTAAAGAC
AAACAATTACATCTAATTAAAATGCTAAGAGATCCTGAGCTGTTAGAGATGAGGAGAGT
AGATAGTATGACCTGATCTTCCCCCCTCTTTTTTTTCCTTTAACAGTATTCTGTTTCAG
CATAAAGCACACTTTCTGAAGAGGTTCCTGGTGGAGACTGGAAATCTGACTGTGTCCTG
TGGCAACACACAGTCCCTTGCATAACTTTGGCTTCAGTCCCTGGATCTGTCCTTTGCAG
CTACGTCAGGTTCCATGGAAGGAGGAAAGAGCTGGAGGGCAGTATCACTCAGCCAAAGC
TCCCATGGGGTCCCATGCTGGCAGGATAA[tc]GGGTTCCTGCTCTAACACAGCTAGCA
CCTCTTCAGGGACATGCTTCCTGTCCACCACCACTTCGTAGACATACTCAGAGAACCAC
TCATCTGTCATGCACAGGTAACCTGGAGAAAAGAACAGAAGACTTATGAGTCCAGAGGG
CAAGGGACAAAGAGCAGAAACCCTTTTTGTAGGATAAACCTTTTACAAAACTAATATTC
ATACATATTTTTCAGCTTTCCCATCTGTAATTTCATTTAATCTAAATCTTATTAGCAAT
TCTGTGAAGCAGATAGGACAGGCATGGCTCTATTTTTAGAAAAATTAGAAAACCGGGTC
TTGAGTAACTAGGTGATGTGCCCAGGTCACATGGTGAGGTTCAGAGCTGGGCCTTGGAC
CTAAGGCTAACACCAGATCCTGTACTGATGCTCTCTTCCTCCGCTGCCTTGGTGATGGT
GAGTGATGACCTGTATACTAG

agaggcacaagaaattaccttgaaaaaatggaattcagtgactgccatctaggaaagac
agtgatactgtccagcagcatgcagttccgagagctcaactcttaggccaccctccctc
cactctactctaggaacaaggagcattaggtctgttttctctccatacacctcaatcgc
tcgtcctctcgtcttattaaaacacagacacagaaccaaactttttgacagttaaagac
aaacaattacatctaattaaaatgctaagagatcctgagctgttagagatgaggagagt
agatagtatgacctgatcttcccccctcttttttttcctttaacagtattctgtttcag
cataaagcacactttctgaagaggttcctggtggagactggaaatctgactgtgtcctg
tggcaacacacagtcccttgcataactttggcttcagtccctggatctgtcctttgcag
ctacgtcaggttccatggaaggaggaaagagctggagggcagtatcactcagccAAAGC
TCCCATGGGGTCCCATGCTGGCAGGATAA[tc]GGGTTCCTGCTCTAACACAGCTAGCA
CCTCTTCAgggacatgcttcctgtccaccaccacttcgtagacatactcagagaaccac
tcatctgtcatgcacaggtaacctggagaaaagaacagaagacttatgagtccagaggg
caagggacaaagagcagaaaccctttttgtaggataaaccttttacaaaactaatattc
atacatatttttcagctttcccatctgtaatttcatttaatctaaatcttattagcaat
tctgtgaagcagataggacaggcatggctctatttttagaaaaattagaaaaccgggtc
ttgagtaactaggtgatgtgcccaggtcacatggtgaggttcagagctgggccttggac
ctaaggctaacaccagatcctgtactgatgctctcttcctccgctgccttggtgatggt
gagtgatgacctgtatactag

FIGURE 23www

53433
CATACGCACATAACACACACACACACACCCACATACAGGCATTGTGAACTAGACACATC
ACCTTACAATCTGTGGTTTACTGGAAGGACATGGAACAAAACCCCCCCAGCCACAGCGT
GGAAGTGCCCTCTCCAGGCACAAGATTCTGCCTCCATGGGGCGTGGTAGCAGCATTGCC
CACCCACCCAGGGCTGAGTGAGCAGGCCTGCCCCACACTGCGCCCATGCACAGCCACTC
CAGGCTGCCTCCCACACTGCCTGCAAGGACCCCAGTGGGGACTGCAAACGGAAGTCTG
CATCCAGGGCCCCAGGGAGGGCAGGTGGGGCTCTGGAGTATAGCACTTTCTAGAAGGGA
AGCACCCTCTTGGTTCTGAACGTAAGTGGGTCTGCTCACAGGGAGGGGCGTGCAGCCAC
CCCAGGACCCCAGCTGTCCAAGGAGCCAGGGAAAACGCACCCACGGGGCACCTACCGCT
GGGAGCGCAAAGAAGGAGATGGCAAAGACAGAGAAGCAGGAGGCGATGGTCTTCCCGAC
CCACGTCTGGGGCACCTTGTCCCCATAGC[ct]GATGGTGGTGACTGTGACCTGCAGGG
AGAGGGACAGTGGTCAGCCACGGATGGGACTGGAGCCTCGGGAGGGCCAACTGCCTAAC
CCAAACCCACCACTCTGATGAGCGGAGAGGCCGGCAAGAGACCCTGACCACCAGGACGA
CCCCGTGTGACTCGGCGAAAGCACCAGGAACAGAGCCGCGGGATGGCACATGTCTCCCA
GGCTCTCGGCGTCACACACAAGGTATGTCCCACCAGCACATGTAAGGAGCCCAGCACCC
ACGAAGGGCCAGGCCTGCTGGCTGGGAACGTGGGCCTGGGAGCTCGCCCCACACCGGCT
GCCTCATCTGCCTGCCTGTCCCCAGGAGGCTGGGCCCCTGGGCCACCGACGTTGCTGTG
CGCCGGCCCCCAGGAGACCGGGAGCTCCCACTGAGGCTGGTCGTCAACAAAGAGCAGGG
GCTGGGATGACGCGCTGCTTC

catacgcacataacacacacacacacacccacatacaggcattgtgaactagacacatc
accttacaatctgtggtttactggaaggacatggaacaaaacccccccagccacagcgt
ggaagtgccctctccaggcacaagattctgcctccatggggcgtggtagcagcattgcc
cacccacccagggctgagtgagcaggcctgccccacactgcgcccatgcacagccactc
caggctgcctcccacactgcctgcaaggaccccagtggggactgcaaacgggaagtctg
catccagggccccagggagggcaggtggggctctggagtatagcactttctagaaggga
agcaccctcttggttctgaacgtaagtgggtctgctcacagggaggggcgtgcagccac
cccaggaccccagctgtccaaggagccagggaaaacgcacccacggggcacctaccgct
gggagcgcaaagaaggagatggcaaagacagagaagcaggaggcgatggtcttcccgaC
CCACGTCTGGGGCACCTTGTCCCCATAGC[ct]GATGGTGGTGACTGTGACCTGCAGGG
AGAGggacagtggtcagccacggatgggactggagcctcgggagggccaactgcctaac
ccaaacccaccactctgatgagcggagaggccggcaagagaccctgaccaccaggacga
ccccgtgtgactcggcgaaagcaccaggaacagagccgcgggatggcacatgtctccca
ggctctcggcgtcacacacaaggtatgtcccaccagcacatgtaaggagcccagcaccc
acgaagggccaggcctgctggctgggaacgtgggcctgggagctcgccccacaccggct
gcctcatctgcctgcctgtccccaggaggctgggcccctgggccaccgacgttgctgtg
cgccggcccccaggagaccgggagctcccactgaggctggtcgtcaacaaagagcaggg
gctgggatgacgcgctgcttc

FIGURE 23xxx

53444
TCAGTTTGTCCAGTAAGATGGGGTGGTCTGTTTCCACCAGGTCCAGCTATCCACTGGTG
GTTCTATGGGGAGCAGTGGGGGTGGTTAAAGGAGCTCTGTGTGGCCGGGAGCGGTGGCT
GATGCCTGTAATCCCAGCTCTTTGGGATGCCAAGGCAGGAGGATCGCTTGAGCCCAGGA
GTTTGAGATCAGGCTGGGCAATATAGTGAAACCTTGTCTCTACGACAAATAAAATTAGC
TAGGCATACTGGTGGTGCACCTGTGGTACCAGCTATAGGGGGGCGCTGAGACAGGAGGA
TTGCTTGAGCTCAGGAGGTTGAGGCTGCAGTGAGCCCTGATTGTGTCACTGCATTCTAG
CCTGGGTGACAGAGTGAGACCCTGTTTAAAAAAAAAAATAGAACTCTGTGTGGCTGAGG
ACAGCTCTCCAGGGGCCCCCACACTGCCTTCCAAATTCCCCTAGGCGGCTACATTGCAC
TAGAAACTATATCCACATCAACCTGTTCACGTCTTTCATGCTGCGAGCTGCGGCCATTC
TCAGCCGAGACCGTCTGCTACCTCGACCT[gt]GCCCCTACCTTGGGGACCAGGCCCTT
GCGCTGTGGAACCAGGTGGGCATCCTCCTTCCGTTCCTCCAAATGGGAATCTTGCTTCT
CTGGTGGGACCAGGAAGTTCTCAGTCCATTTCCTATCTCCTACACTCTCCACAGTTTAT
CTGAGTTGGGAGGGTCCCTCTCCAAATGTGTCTTGGGGTGGGGGATCAAGACACATTTG
GAGAGGGAACCTCCCAACTCGGCCTCTGCCATCATTTAACTCTCCCAGCCTATCACTCC
CATACTGGAATTTTCCGTTCCTCTCCCTCATTATTTCACCCATCATTGAACTTTTTCAC
CAATGAGAGAATCCACCTGCTGGCGGTGAGGCATGGCAGGATACGAGAAAGTAAGTGGG
GGTGGGGATGTGGCAGGTGCCAGTTTGTTACTAGGAGACAGGGTGGGAGAGACTAGAGT
CTGGGAGCAGACGTGGTAAGA

tcagtttgtccagtaagatggggtggtctgtttccaccaggtccagctatccactggtg
gttctatggggagcagtgggggtggttaaaggagctctgtgtggccgggagcggtggct
gatgcctgtaatcccagctctttgggatgccaaggcaggaggatcgcttgagcccagga
gtttgagatcaggctgggcaatatagtgaaaccttgtctctacgacaaataaaattagc
taggcatactggtggtgcacctgtggtaccagctataggggggcgctgagacaggagga
ttgcttgagctcaggaggttgaggctgcagtgagccctgattgtgtcactgcattctag
cctgggtgacagagtgagaccctgtttaaaaaaaaaaatagaactctgtgtggctgagg
acagctctccaggggcccccacactgccttccaaattcccctaggcggctacattgcac
tagaaactatatccacatcaacctgttcacgtctttcatgctgcgagctgcggccattc
tcagccGAGACCGTCTGCTACCTCGACCT[gt]GCCCCTACCTTGGGGACCAGGCCctt
gcgctgtggaaccaggtgggcatcctccttccgttcctccaaatgggaatcttgcttct
ctggtgggaccaggaagttctcagtccatttcctatctcctacactctccacagtttat
ctgagttgggagggtccctctccaaatgtgtcttggggtgggggatcaagacacatttg
gagagggaacctcccaactcggcctctgccatcatttaactctcccagcctatcactcc
catactggaattttccgttcctctccctcattatttcacccatcattgaactttttcac
caatgagagaatccacctgctggcggtgaggcatggcaggatacgagaaagtaagtggg
ggtggggatgtggcaggtgccagtttgttactaggagacagggtgggagagactagagt
ctgggagcagacgtggtaaga

FIGURE 23yyy

53445
TGTTTCCACCAGGTCCAGCTATCCACTGGTGGTTCTATGGGGAGCAGTGGGGGTGGTTA
AAGGAGCTCTGTGTGGCCGGGAGCGGTGGCTGATGCCTGTAATCCCAGCTCTTTGGGAT
GCCAAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGATCAGGCTGGCAATATAGTG
AAACCTTGTCTCTACGACAAATAAAATTAGCTAGGCATACTGGTGGTGCACCTGTGGTA
CCAGCTATAGGGGGGCGCTGAGACAGGAGGATTGCTTGAGCTCAGGAGGTTGAGGCTGC
AGTGAGCCCTGATTGTGTCACTGCATTCTAGCCTGGGTGACAGAGTGAGACCCTGTTTA
AAAAAAAAAATAGAACTCTGTGTGGCTGAGGACAGCTCTCCAGGGGCCCCCACACTGCC
TTCCAAATTCCCCTAGGCGGCTACATTGCACTAGAAACTATATCCACATCAACCTGTTC
ACGTCTTTCATGCTGCGAGCTGCGGCCATTCTCAGCCGAGACCGTCTGCTACCTCGACC
TGGCCCCTACCTTGGGGACCAGGCCCTTG[ct]GCTGTGGAACCAGGTGGGCATCCTCC
TTCCGTTCCTCCAAATGGGAATCTTGCTTCTCTGGTGGGACCAGGAAGTTCTCAGTCCA
TTTCCTATCTCCTACACTCTCCACAGTTTATCTGAGTTGGGAGGGTCCCTCTCCAAATG
TGTCTTGGGGTGGGGGATCAAGACACATTTGGAGAGGGAACCTCCCAACTCGGCCTCTG
CCATCATTTAACTCTCCCAGCCTATCACTCCCATACTGGAATTTTCCGTTCCTCTCCCT
CATTATTTCACCCATCATTGAACTTTTTCACCAATGAGAGAATCCACCTGCTGGCGGTG
AGGCATGGCAGGATACGAGAAAGTAAGTGGGGGTGGGGATGTGGCAGGTGCCAGTTTGT
TACTAGGAGACAGGGTGGGAGAGACTAGAGTCTGGGAGCAGACGTGGTAAGAACTAACT
TGTTGAAAGTTGGACCATACC

tgtttccaccaggtccagctatccactggtggttctatggggagcagtgggggtggtta
aaggagctctgtgtggccgggagcggtggctgatgcctgtaatcccagctctttgggat
gccaaggcaggaggatcgcttgagcccaggagtttgagatcaggctggcaatatagtg
aaaccttgtctctacgacaaataaaattagctaggcatactggtggtgcacctgtggta
ccagctatagggggcgctgagacaggaggattgcttgagctcaggaggttgaggctgc
agtgagccctgattgtgtcactgcattctagcctgggtgacagagtgagaccctgttta
aaaaaaaaaatagaactctgtgtggctgaggacagctctccaggggcccccacactgcc
ttccaaattcccctaggcggctacattgcactagaaactatatccacatcaacctgttc
acgtctttcatgctgcgagctgcggccattctcagccgagaccgtctgctacctcgacc
tggCCCCTACCTTGGGGACCAGGCCCTTG[ct]GCTGTGGAACCAGGTGGGCATCCTCC
ttccgttcctccaaatgggaatcttgcttctctggtgggaccaggaagttctcagtcca
tttcctatctcctacactctccacagtttatctgagttgggagggtccctctccaaatg
tgtcttggggtgggggatcaagacacatttggagagggaacctcccaactcggcctctg
ccatcatttaactctcccagcctatcactcccatactggaattttccgttcctctccct
cattatttcacccatcattgaactttttcaccaatgagagaatccacctgctggcggtg
aggcatggcaggatacgagaaagtaagtgggggtggggatgtggcaggtgccagtttgt
tactaggagacagggtgggagagactagagtctgggagcagacgtggtaagaactaact
tgttgaaagttggaccatacc

FIGURE 23zzz

53459
CCTTTTATTTTTCTTCCATGGAATTTTCCAGTTAACTTGAGAAAGTGGAATCGAATTCC
GATGTTGAATTTTCCTTCTGGCCCCATTCATGTGGCAGGTGGTGATTCAGGTACTACTG
GGGGCTGCTCAGACAAACCTCCTCATCAGACATCAAGAGGCTGTTGCACCAGGAGGGCC
GGTACCGTGTCTAGAGGTGGTCGGCATGGGGTTGGAGTTGTATTACATAAACCCTACTC
CAAACAAATGCATGGGGATGTGGCTGGAGTTCCCCGTTGTCTAACCAGTGCCAAAGGGC
AGGACGGTACCTCACCCCACGTTCTTAACTATGGGTTGGCAACATGTTCCTGGATGTGT
TTGCTGGCACAGTGACAGGTGCTAGCAACCAGGGTGTTGACACAGTCCAACTCCATCCT
CACCAGGTCACTGGCTGGAACCCCTGGGGGCCACCATTGCGGGAATCAGCCTTTGAAAC
GATGGCCAACAGCAGCTAATAATAAACCAGTAATTTGGGATAGACGAGTAGCAAGAGGG
CATTGGTTGGTGGGTCACCCTCCTTCTCA[gt]AACACATTATAAAAACCTTCCGTTTC
CACAGGATTGTCTCCCGGGCTGGCAGCAGGGCCCCAGCGGCACCATGTCTGCCCTCGGA
GTCACCGTGGCCCTGCTGGTGTGGGCGGCCTTCCTCCTGCTGGTGTCCATGTGGAGGCA
GGTGCACAGCAGCTGGAATCTGCCCCCAGGCCCTTTCCCGCTTCCCATCATCGGGAACC
TCTTCCAGTTGGAATTGAAGAATATTCCCAAGTCCTTCACCCGGGTAAGAGAAATAGTG
TTGATTTTAGGGAGAATAACTCAGCAATTGGATCTGGTATGTGTGTATTCAACTCATTT
GCAGACAAATTGTGGTTGTTCAATACCAGCCTGTTGTGAATTACCTGAATTGATAGCAT
CCTGGAGCGACACTCAAAATGTGTCGCCTGTGGTGCAGCTGGAGCCCGGAGCCTGCGTG
CCAGGCCCCGGAGGCCCCCGC

cctttatttttcttccatggaattttccagttaacttgagaaagtggaatcgaattcc
gatgttgaattttccttctggccccattcatgtggcaggtggtgattcaggtactactg
ggggctgctcagacaaacctcctcatcagacatcaagaggctgttgcaccaggagggcc
ggtaccgtgtctagaggtggtcggcatggggttggagttgtattacataaaccctactc
caaacaaatgcatggggatgtggctggagttccccgttgtctaaccagtgccaaagggc
aggacggtacctcaccccacgttcttaactatgggttggcaacatgttcctggatgtgt
ttgctggcacagtgacaggtgctagcaaccagggtgttgacacagtccaactccatcct
caccaggtcactggctggaacccctgggggccaccattgcgggaatcagcctttgaaac
gatggccaacagcagctaataataaaccagtaatttgggatagacGAGTAGCAAGAGGG
CATTGGTTGGTGGGTCACCCTCCTTCTCA[gt]AACACATTATAAAAACCTTCCGTTTC
CACAGGATTGTCTCCCGggctggcagcagggccccagcggcaccatgtctgccctcgga
gtcaccgtggccctgctggtgtgggcggccttcctcctgctggtgtccatgtggaggca
ggtgcacagcagctggaatctgcccccaggccctttcccgcttcccatcatcgggaacc
tcttccagttggaattgaagaatattcccaagtccttcacccgggtaagagaaatagtg
ttgattttagggagaataactcagcaattggatctggtatgtgtgtattcaactcattt
gcagacaaattgtggttgttcaataccagcctgttgtgaattacctgaattgatagcat
cctggagcgacactcaaaatgtgtcgcctgtggtgcagctggagcccggagcctgcgtg
ccaggccccggaggcccccgc

FIGURE 23aaaa

53464
AAGCAGTACCAGCAGCCAGAAACCGCATAACAAATACATTGGGCAATTGGGAGTTGGGG
ATGTTGACTGAACCTTTGAACTTGACTGATCCGAATGCCAAATTCTAATTTAAAAAGGG
AAAACTGGATGTTTGAGACATAGTGTGATTTGGTGAAGCAAACAATGGACTCCCAGGTT
AAAAACTCTAATTAGTCTCTTCTGCTGAGTTCCTGAGTTAACCGTTTGTGTAGTGGTCT
CCTAGTATATTTTATAACTTACAAAGCTAGAGGATCAAAGCAATTATCTAGAAATACAC
ACAAAACTCATGTTTGGTATAAATGTCTGAACAATTAACCAAACTGTCGCAACAGCTTT
TTTCATTGATTTGATCTAACACTGATATGCCTCATAGGGTCATGAGTTGAAAAAACAAC
TCTAAAGCTATTCCACAAGAAGAAAGATAATATTTTTCTAAACAAGTGTTAGGAAAATG
AAAATATGAAAGTTTCTGTTTATGCTTATTTATGAAATTTGCCTACCTTCCAAGTGTGT
CCCCAAGCCACCCACCAAAGAATGATGCA[ag]TCATTCCACCAACTGCAAAGCTGGAT
ACAGACAGGGACCAGAGCATGGTGATTAGTTGAGCAGCTGCCACAGTCTCTTCCTCAGC
CCAAGGGGTTGGTTTTGGGTTCATTGAGTATGAGATTGTGGCAGTTCATCTGTACTGT
TGATAACATAGTTGTTGATAGCTTTTCGGTCATCCAGTGGAACACCCAAAACATGTCTA
TAGTGAGATATTATTACCTAGGAGATAAAGAAAAATAGCTTTACTATTTCAAACATTCT
ATGTATTTTTGTTTTTGTCTTTAAAGTGTTTGTTACGTGTTTAAATAGTACCATCTCAA
TTATGTGTTTTATATACATATAAACATGGATAGATTTGTTTACAGTTGGCCATATCCTA
TAAAAGAAAGGTTATAAATTACATTGCCAACAAGAACCAGGCAGGAACAAATAAATGAA
GGGAACATGTAATACTTTGAT

aagcagtaccagcagccagaaaccgcataacaaatacattgggcaattgggagttgggg
atgttgactgaacctttgaacttgactgatccgaatgccaaattctaatttaaaaaggg
aaaactggatgtttgagacatagtgtgatttggtgaagcaaacaatggactcccaggtt
aaaaactctaattagtctcttctgctgagttcctgagttaaccgtttgtgtagtggtct
cctagtatattttataacttacaaagctagaggatcaaagcaattatctagaaatacac
acaaaactcatgtttggtataaatgtctgaacaattaaccaaactgtcgcaacagcttt
tttcattgatttgatctaacactgatatgcctcatagggtcatgagttgaaaaaacaac
tctaaagctattccacaagaagaaagataatatttttctaaacaagtgttaggaaaatg
aaaatatgaaagtttctgtttatgcttatttatgaaatttgcctaccttccaaGTGTGT
CCCCAAGCCACCCACCAAAGAATGATGCA[ag]TCATTCCACCAACTGCAAAGCTGGAT
ACAGACAGGgaccagagcatggtgattagttgagcagctgccacagtctcttcctcagc
ccaaggggttggttttgggttcattgagtatgagattgtggcagttcatctgtactgt
tgataacatagttgttgatagcttttcggtcatccagtggaacacccaaaacatgtcta
tagtgagatattattacctaggagataaagaaaaatagctttactatttcaaacattct
atgtatttttgtttttgtctttaaagtgtttgttacgtgtttaaatagtaccatctcaa
ttatgtgttttatatacatataaacatggatagatttgtttacagttggccatatccta
taaaagaaaggttataaattacattgccaacaagaaccaggcaggaacaaataaatgaa
gggaacatgtaatactttgat

FIGURE 23bbbb

53470
ATGCCCTTTGGCCTAAACCCTGGACTTGACTAAGAAATGCAGCCTCCAATGACATTGCG
GGAAAAGGGAATCTGGGAACTTCTATGACACAATTCAGTCTTGCTGAGCATTTGGGGCT
AATATTTAACTCTGAACATATATTGACATAGGCAATTCTTCCATAACAGATTCATACAA
AATTTAAAAATGCATATAGAAGCCTTAATTTTTATTTAAATTCTTTTATTTAATTGTGT
TTTAGAGGCAGAGAATAGTGTGTCTTTTTTTGCCTCTTTTATAATTTTTATTTTTTTTT
TTCATTTTTGCCACTGTCTTTCTTTGCGCTTTCTAGGGCATTACATTTTTCTTTTCCGT
TTTCTCCATGTTTCTTAGCGAGATTCTCTAAAAGGTTACTTCTATTTCCATCACATCAT
CATCTAGCTCCAGCAGGCCTACTTTTCTTCATTTCCTCTATTGTATTTTCTGCTTTTCA
TTCTTGCTGTCTGCTCCTCTCTCATCATCCTTGCCTCTGTCTGTTTAATCCTCCTGTCC
TTCATTTTCCTTTTTTGCCTCTGCATTCA[gt]CATTTCTACTTCCAATCTCCCTCCTC
TGCTCTTTCTTCTTTCCTCTGATCTGCAGACTTGCTTCTGTCCCCTCCTTCTGTTCCCC
TCCTGGATGTGTCTTTGGCCAACCTTTCCTTCTCTGAGACTTCGTGTTCTTGTTGGTAG
ATGGGGGCTGATACTGTAAACATCACAAAAATAATTGCATTGAGAACAAGTGGTTCCCA
TGGTGTCCCTTTGAATGAGCTCAGAATGCCCAGGCTCCATATGATGCAGGAGACAGCAC
TCATGCTGGAGAGGGGTCTAGACCTCAGTCACAAGACCCACCATTCCAGAACTTTGGGA
CTCATCTCTTGACACCTACCCCCTCCCCAGTTAGAAACCAAGAGGCGCTGGGTCACCTG
GGAAGAGAAAGAATGAATCTGCCTTTGCCCCAGCAAGCACGCTTTCCTGCCACATTCAC
CTAAAAGTCTTTTCTGAGATC

atgccctttggcctaaaccctggacttgactaagaaatgcagcctccaatgacattgcg
ggaaaagggaatctgggaacttctatgacacaattcagtcttgctgagcatttggggct
aatatttaactctgaacatatattgacataggcaattcttccataacagattcatacaa
aatttaaaaatgcatatagaagccttaatttttatttaaattcttttatttaattgtgt
tttagaggcagagaatagtgtgtctttttttgcctcttttataatttttattttttttt
ttcatttttgccactgtctttctttgcgctttctagggcattacatttttcttttccgt
tttctccatgtttcttagcgagattctctaaaaggttacttctatttccatcacatcat
catctagctccagcaggcctacttttcttcatttcctctattgtattttctgcttttca
ttcttgctgtctgctcctctctcatcatccttgcctctgtctgTTTAATCCTCCTGTCC
TTCATTTTCCTTTTTTGCCTCTGCATTCA[gt]CATTTCTACTTCCAATCTCCCTCCTC
TGCTCTTTCTTCTTTCCTCtgatctgcagacttgcttctgtcccctccttctgttcccc
tcctggatgtgtctttggccaacctttccttctctgagacttcgtgttcttgttggtag
atgggggctgatactgtaaacatcacaaaaataattgcattgagaacaagtggttccca
tggtgtccctttgaatgagctcagaatgcccaggctccatatgatgcaggagacagcac
tcatgctggagaggggtctagacctcagtcacaagacccaccattccagaactttggga
ctcatctcttgacacctaccccctccccagttagaaaccaagaggcgctgggtcacctg
ggaagagaaagaatgaatctgcctttgccccagcaagcacgctttcctgccacattcac
ctaaaagtcttttctgagatc

FIGURE 23cccc

53475
CCATAGCGAATGTTTTCAGCTATCGTGGTGGCAAACAATACAGGTTCCTGACTCACCAC
ACCAATGATTTCCCGTAGAAACCTTACATTTATGGTCCTAATATCCTGTCCATCAACAC
TGACCTGGAATAAAAAGTAAGTGTGACTTTCATACATTTGTAATTGAAAGGGCAACATC
AGAAAGATGTGCAATGTGACTGCTGATCACCGCAGGGTCTAGCTCGCATGGGTCATCTC
ACCATCCCCTCTGTGGGGTCATAGAGCCTCTGCATCAGCTGGACTGTTGTGCTCTTCCC
ACAGCCACTGTTTCCAACCAGGGCCACCGTCTGCCCACTCTGCACCTTCAGGTTCAGAC
CCTTCAAGATCTACCAGGACGAGTGAGAAAAAAACTTCAAGGCAATTCACAGACACAGG
ATATAGGAACTGACTGTTCACTAGGTTTAAATATACATGCACTTTTTTATAATCTCTAC
AAGAAAACATCAGAAACTCTTCATTCAATAGATTAATTGTTGATTAATCATTTATCACT
GTACCTTAACTTCTTTTCGAGATGGGTAA[ct]TGAAGTGAACATTTCTGAATTCCAAA
TTTCCCTTAATATTATCTGGTTTGTGCCCACTCTTCGAATAGCTGTCAATACTTGGCTT
CTAAACAGAATCAAATTTTAAGAGATTACTAGGTTACAATAACTACTTTTAGTGATATT
TTGTGGAGAGCTGGATAAAGTGACAAAGAAATTGACTTAACTGGACAATCTTTTAGATA
GGTGGATAGATGGCCAACTCAGACTTACATTATCAATTATCTTGAAGATTTCATAAGCT
GCTCCTCTTGCATTTGCAAATGCTTCAATGCTTGGAGATGCCTGTCCAACACTAAAAGC
CCCAATTAATACAGAAAAGAATACCTGAGGAATGTGAAGAAAAACCATCAGGCTACTGA
GATAGTGACAGCAATTTTTTTCATACTTCTTCTGTCTTTTTCTAACATAGGTAATTAA
AATTTAAAATGGCGAGGCAAC

ccatagcgaatgttttcagctatcgtggtggcaaacaatacaggttcctgactcaccac
accaatgatttcccgtagaaaccttacatttatggtcctaatatcctgtccatcaacac
tgacctggaataaaaagtaagtgtgactttcatacatttgtaattgaaagggcaacatc
agaaagatgtgcaatgtgactgctgatcaccgcagggtctagctcgcatgggtcatctc
accatcccctctgtggggtcatagagcctctgcatcagctggactgttgtgctcttccc
acagccactgtttccaaccagggccaccgtctgcccactctgcaccttcaggttcagac
ccttcaagatctaccaggacgagtgagaaaaaaacttcaaggcaattcacagacacagg
atataggaactgactgttcactaggtttaaatatacatgcacttttttataatctctac
aagaaaacatcagaaactcttcattcaatagatTAATTGTTGATTAATCATTTATCACT
GTACCTTAACTTCTTTTCGAGATGGGTAA[ct]TGAAGTGAACATTTCTGAATTCCAAA
TTTCCCTTAATATTATCTGGTTTGTGCCCactcttcgaatagctgtcaatacttggctt
ctaaacagaatcaaattttaagagattactaggttacaataactacttttagtgatatt
ttgtggagagctggataaagtgacaaagaaattgacttaactggacaatcttttagata
ggtggatagatggccaactcagacttacattatcaattatcttgaagatttcataagct
gctcctcttgcatttgcaaatgcttcaatgcttggagatgcctgtccaacactaaaagc
cccaattaatacagaaaagaatacctgaggaatgtgaagaaaaaccatcaggctactga
gatagtgacagcaatttttttcatacttcttctgtctttttctaacataggtaattaa
aatttaaaatggcgaggcaac

FIGURE 23dddd

53478
GGACAAGAGCAGGGCTTTAAATGCCCCATAAATATGTGTGGCAAGGATGAAAGCACATA
GGACTCAAAGAGGAACAAAGGAGCAGAAAGGCAGGAAGAGTTGGTGCTGCCTTCAAAGG
AGAGTAGGAACGAGGGCAGGTGGTATCAGGTGGACCTCTATGTGGTCCTGGGTTACAAA
GGTGCCAGGAAAAAGCAAGAAATGGAAGAGTCTAAAAAGCAATGGAAGATTGTGGAAAA
TGATGGAAGATTCCGGAAAGTGGTGGAAGATTCCAGAAAATGATGGAAGATTCCAGAAA
GTGATGAAAGATTCTGGAAAGCAATGAAACATTCCAGAAAGTGATGAGACAGTGATAGA
GTCTGGTTCCAGGCGAAGTGGGAGAGGATGGGATTTGAGAAGGGAATGATCCCTCCTCA
CACCTCTAGGATGGGAAGCTTAGTGGAGTGAGGGGTGGGTAGGAGGTTACACCCTGTGT
CCTCTGTCGCTCTGTGCAGGAGGAGGAGGCAGAGAAAGGGAAGGGTCAGGAAAGCCAGC
CCATGTCCCACCCCCACTGGACTCACCAC[ga]TGATGGCAGGTGAAGCCCTTCATGAC
CGAGGCCTCATTGAGGAACTCAATCCGCTCTCGGAGACTGGCTGACTCGTTGACCGTCT
TCACCGCCACGCGGGTCTCTGCCTCACCCTTGATGATGTCCCTGGCATTGCCCTCATAC
ACCATGCCGAAGGAGCCCTGCCCCAGCTCTCGAAGGAGGGTGATCTTCTCTCGAGACAC
CTCCCACTCGTCCGGCACGTACACAGAGCATGGAAACACTACTTCTTACTTATCTACAC
AGCATCCTTGGAGGATCCCTTGGGGGTCTGCAGCCACCTTCCACCCAAGCCCTCACCCA
AACCCCCTCGAAAACACTCATGAAATGAGTTCTGTGATCCAGGACCCATGCCGGGCACT
GGGCATATGGCCGAGAACAGGACAGGCATCTGCACCCATGGAGAGGGCATGGCAGAGAC
TCAAGGAAGGAGCCACAACTG

ggacaagagcagggctttaaatgccccataaatatgtgtggcaaggatgaaagcacata
ggactcaaagaggaacaaaggagcagaaaggcaggaagagttggtgctgccttcaaagg
agagtaggaacgagggcaggtggtatcaggtggacctctatgtggtcctgggttacaaa
ggtgccaggaaaaagcaagaaatggaagagtctaaaaagcaatggaagattgtggaaaa
tgatggaagattccggaaagtggtggaagattccagaaaatgatggaagattccagaaa
gtgatgaaagattctggaaagcaatgaaacattccagaaagtgatgagacagtgataga
gtctggttccaggcgaagtgggagaggatgggatttgagaagggaatgatccctcctca
cacctctaggatgggaagcttagtggagtgaggggtgggtaggaggttacaccctgtgt
cctctgtcgctctgtgcaggaggaggaggcagagaaagggaagggtcaggaaagccagc
CCATGTCCCACCCCCACTGGACTCACCAC[ga]TGATGGCAGGTGAAGCCCTTCATGAC
CGAggcctcattgaggaactcaatccgctctcggagactggctgactcgttgaccgtct
tcaccgccacgcgggtctctgcctcacccttgatgatgtccctggcattgccctcatac
accatgccgaaggagccctgccccagctctcgaaggagggtgatcttctctcgagacac
ctcccactcgtccggcacgtacacagagcatggaaacactacttcttacttatctacac
agcatccttggaggatcccttgggggtctgcagccaccttccacccaagccctcaccca
aacccccctcgaaaacactcatgaaatgagttctgtgatccaggacccatgccgggcact
gggcatatggccgagaacaggacaggcatctgcacccatggagagggcatggcagagac
tcaaggaaggagccacaactg

FIGURE 23eeee

53488
TCCCACCTCCTGGGCAGCCTGGTAGAGGAGACATTCCTTTAATTCTTCCTGCCTAATTT
AGAGGCTGGGTGGGGGTCTGAAGGTTCACTCCCTTCACATCATCCCACTAGTCTACTTT
GGGAAGAATTACAGGTTGTTGGAGCTGGAAGCCCCATTCTAGGCATGGTCTGAAGACCT
GAACAATCCCAGGGGGTGGTGAAGGGGGCAGGGAGGAGATGGGCACCACTTACCATTTG
AGGCCGCCCAGAGAAGTCCTTGCCCTTCTTAATAAGCACCTCCTTGGCCAGCTGGTGGT
GGCCGACAATCACTGTAGTCTTGGTGCCCATACGAACCGAATAGATGGGGCCATATTTT
TTCTGCAGCTTGAAGAAGTTGTTATGCATATGGCCGTGTCTGGGGAGGAATGGCAGGCT
GCCCACCAGGGGCAGGGACAGGAGGCTCTTGGGGTACTTGGCACCAGGGCACCTTCTCT
TGGGCCAAAACAAATAAGCTAGGGTAAGCAGCAAGAGAGCCACGAGCTCCCACATGGTG
GCTGGGTGCCGGCAGGCAAGATAGACAGC[ag]GTGGAGTAGAAGAGCTGTGGCAACTC
TAGGGCACAAGGAGGCCTTTTAAAGGGCTACCCTGATCTTCACCTTGACTTTGTGTTAT
CTCTTGCCTTGTGGAAAGATTCTCCTGGAGCCCAGCCAGGCCTGAGCTCATATCCAGAA
GGGAGAGAGGCGGTGGGAGTGAAGGCCTCCTCAAGGGCTGGCTCAACTCCAGGGCAAAC
CTCCGGAGGAGGAGCTAGGTAAGGGAGGTCAGTTGATCACCCTCTGAGGAGCTCCCCAT
GCTTGAATGACTCCAGAGTGCGAATGGTATCTGGGCTCAGGAGTCAAGGCTTGGAACTT
TCCATGTTGCAAAATCAAAATCACTGGACAGATGACAGATTCAGGAGGGTCACAAGTAG
CAGGGACTGTTAAAGGTCTTTTATGCTTCTTTTTTTTTTTTTCAGAGTCTTGCTCCATC
ACCAGGCTGGTGTGCAGTGGT

tcccacctcctgggcagcctggtagaggagacattcctttaattcttcctgcctaattt
agaggctgggtgggggtctgaaggttcactcccttcacatcatcccactagtctacttt
gggaagaattacaggttgttggagctggaagccccattctaggcatggtctgaagacct
gaacaatcccagggggtggtgaagggggcagggaggagatgggcaccacttaccatttg
aggccgcccagagaagtccttgcccttcttaataagcacctccttggccagctggtggt
ggccgacaatcactgtagtcttggtgcccatacgaaccgaatagatggggccatatttt
ttctgcagcttgaagaagttgttatgcatatggccgtgtctggggaggaatggcaggct
gcccaccaggggcagggacaggaggctcttggggtacttggcaccagggcaccttctct
tgggccaaaacaaataagctagggtaagcagcaagagagccacgagctcccacatGGTG
GCTGGGTGCCGGCAGGCAAGATAGACAGC[ag]GTGGAGTAGAAGAGCTGTGGCAACTC
TAGGGCAcaaggaggccttttaaagggctaccctgatcttcaccttgactttgtgttat
ctcttgccttgtggaaagattctcctggagcccagccaggcctgagctcatatccagaa
gggagagaggcggtgggagtgaaggcctcctcaagggctggctcaactccagggcaaac
ctccggaggaggagctaggtaagggaggtcagttgatcaccctctgaggagctccccat
gcttgaatgactccagagtgcgaatggtatctgggctcaggagtcaaggcttggaactt
tccatgttgcaaaatcaaaatcactggacagatgacagattcaggagggtcacaagtag
cagggactgttaaaggtcttttatgcttctttttttttttttcagagtcttgctccatc
accaggctggtgtgcagtggt

FIGURE 23ffff

53489
CAATAGCTAGGCTAATTCTCCCCAGCAGCTTTCATGGAGGACAGTAGTCACTGCCCCCA
TTTTCCATGAAAAGTAACATGAATCCTGGCTGTATAAGGGGCACTTACTGTGCTGGGTG
CTAGGCTAAGTGCTGTACATGCACCTTCTCAGTCCATTAGAGAAGTCTAGGCTCAGAGA
GAGGAGTGGAGTGAGGATTCCTTGACCCCTCAGACCACTGTGGTCCTCCCATCCCACCT
CCTGGGCAGCCTGGTAGAGGAGACATTCCTTTAATTCTTCCTGCCTAATTTAGAGGCTG
GGTGGGGGTCTGAAGGTTCACTCCCTTCACATCATCCCACTAGTCTACTTTGGGAAGAA
TTACAGGTTGTTGGAGCTGGAAGCCCCATTCTAGGCATGGTCTGAAGACCTGAACAATC
CCAGGGGGTGGTGAAGGGGGCAGGGAGGAGATGGGCACCACTTACCATTTGAGGCCGCC
CAGAGAAGTCCTTGCCCTTCTTAATAAGCACCTCCTTGGCCAGCTGGTGGTGGCCGACA
ATCACTGTAGTCTTGGTGCCCATACGAAC[ca]GAATAGATGGGGCCATATTTTTTCTG
CAGCTTGAAGAAGTTGTTATGCATATGGCCGTGTCTGGGGAGGAATGGCAGGCTGCCCA
CCAGGGGCAGGGACAGGAGGCTCTTGGGGTACTTGGCACCAGGGCACCTTCTCTTGGGC
CAAAACAAATAAGCTAGGGTAAGCAGCAAGAGAGCCACGAGCTCCCACATGGTGGCTGG
GTGCCGGCAGGCAAGATAGACAGCGGTGGAGTAGAAGAGCTGTGGCAACTCTAGGGCAC
AAGGAGGCCTTTTAAAGGGCTACCCTGATCTTCACCTTGACTTTGTGTTATCTCTTGCC
TTGTGGAAAGATTCTCCTGGAGCCCAGCCAGGCCTGAGCTCATATCCAGAAGGGAGAGA
GGCGGTGGGAGTGAAGGCCTCCTCAAGGGCTGGCTCAACTCCAGGGCAAACCTCCGGAG
GAGGAGCTAGGTAAGGGAGGT

caatagctaggctaattctccccagcagctttcatggaggacagtagtcactgccccca
ttttccatgaaaagtaacatgaatcctggctgtataaggggcacttactgtgctgggtg
ctaggctaagtgctgtacatgcaccttctcagtccattagagaagtctaggctcagaga
gaggagtggagtgaggattccttgacccctcagaccactgtggtcctcccatcccacct
cctgggcagcctggtagaggagacattcctttaattcttcctgcctaatttagaggctg
ggtgggggtctgaaggttcactcccttcacatcatcccactagtctactttgggaagaa
ttacaggttgttggagctggaagccccattctaggcatggtctgaagacctgaacaatc
ccagggggtggtgaagggggcagggaggagatgggcaccacttaccatttgaggccgcc
cagagaagtccttgcccttcttaataagcacctccttggccagctggtggtGGCCGACA
ATCACTGTAGTCTTGGTGCCCATACGAAC[ca]GAATAGATGGGGCCATATTTTTTCTG
CAGCTTGAAGAagttgttatgcatatggccgtgtctggggaggaatggcaggctgccca
ccaggggcagggacaggaggctcttggggtacttggcaccagggcaccttctcttgggc
caaaacaaataagctagggtaagcagcaagagagccacgagctcccacatggtggctgg
gtgccggcaggcaagatagacagcggtggagtagaagagctgtggcaactctagggcac
aaggaggccttttaaagggctaccctgatcttcaccttgactttgtgttatctcttgcc
ttgtggaaagattctcctggagcccagccaggcctgagctcatatccagaagggagaga
ggcggtgggagtgaaggcctcctcaagggctggctcaactccagggcaaacctccggag
gaggagctaggtaagggaggt

FIGURE 23gggg

53503
GGGTTTCCTGTTTCCTTTTCTGATCATTCTTACAAGTTATACTCTTATTTGGAAGGCCC
TAAAGAAGGCTTATGAAATTCAGAAGAACAAACCAAGAAATGATGATATTTTTAAGATA
ATTATGGCAATTGTGCTTTTCTTTTTCTTTTCCTGGATTCCCCACCAAATATTCACTTT
TCTGGATGTATTGATTCAACTAGGCATCATACGTGACTGTAGAATTGCAGATATTGTGG
ACACGGCCATGCCTATCACCATTTGTATAGCTTATTTTAACAATTGCCTGAATCCTCTT
TTTTATGGCTTTCTGGGGAAAAAAATTTAAAAGATATTTTCTCCAGCTTCTAAAATATAT
TCCCCCAAAAGCCAAATCCCACTCAAACCTTTCAACAAAAATGAGCACGCTTTCCTACC
GCCCCTCAGATAATGTAAGCTCATCCACCAAGAAGCCTGCACCATGTTTTGAGGTTGAG
TGACATGTTCGAAACCTGTCCATAAAGTAATTTTGTGAAAGAAGGAGCAAGAGAACATT
CCTCTGCAGCACTTCACTACCAAATGAGC[ac]TTAGCTACTTTTCAGAATTGAAGGAG
AAAATGCATTATGTGGACTGAACCGACTTTTCTAAAGCTCTGAACAAAAGCTTTTCTTT
CCTTTTGCAACAAGACAAAGCAAAGCCACATTTTGCATTAGACAGATGACGGCTGCTCG
AAGAACAATGTCAGAAACTCGATGAATGTGTTGATTTGAGAAATTTTACTGACAGAAAT
GCAATCTCCCTAGCCTGCTTTTGTCCTGTTATTTTTTATTTCCACATAAAGGTATTTAG
AATATATTAAATCGTTAGAGGAGCAACAGGAGATGAGAGTTCCAGATTGTTCTGTCCAG
TTTCCAAAGGGCAGTAAAGTTTTCGTGCCGGTTTTCAGCTATTAGCAACTGTGCTACAC
TTGCACCTGGTACTGCACATTTTGTACAAAGATATGCTAAGCAGTAGTCGTCAAGTTGC
AGATCTTTTTGTGAAATTCAA

gggtttcctgtttccttttctgatcattcttacaagttatactcttatttggaaggccc
taaagaaggcttatgaaattcagaagaacaaaccaagaaatgatgatatttttaagata
attatggcaattgtgcttttctttttcttttcctggattccccaccaaatattcacttt
tctggatgtattgattcaactaggcatcatacgtgactgtagaattgcagatattgtgg
acacggccatgcctatcaccatttgtatagcttattttaacaattgcctgaatcctctt
ttttatggctttctggggaaaaaaatttaaaagatattttctccagcttctaaaatatat
tcccccaaaagccaaatcccactcaaacctttcaacaaaaatgagcacgctttcctacc
gcccctcagataatgtaagctcatccaccaagaagcctgcaccatgttttgaggttgag
tgacatgttcgaaacctgtccataaagtaattttgtgAAAGAAGGAGCAAGAGAACATT
CCTCTGCAGCACTTCACTACCAAATGAGC[ac]TTAGCTACTTTTCAGAATTGAAGGAG
AAAATGCATTATGTGGACTGAACCGacttttctaaagctctgaacaaaagcttttcttt
ccttttgcaacaagacaaagcaaagccacattttgcattagacagatgacggctgctcg
aagaacaatgtcagaaactcgatgaatgtgttgatttgagaaattttactgacagaaat
gcaatctccctagcctgcttttgtcctgttattttttatttccacataaaggtatttag
aatatattaaatcgttagaggagcaacaggagatgagagttccagattgttctgtccag
tttccaaagggcagtaaagttttcgtgccggttttcagctattagcaactgtgctacac
ttgcacctggtactgcacattttgtacaaagatatgctaagcagtagtcgtcaagttgc
agatctttttgtgaaattcaa

FIGURE 23hhhh

53517
GGGAGAGAGGACCTGTGACAGGATAAAGGGGCTGCCTTATTTAAACCTGGAAGGAAGAA
CGACAGTATAAGCTTCCAGGATATTAATATCAGGCTAACATGGACAGTTAAGAGCCTTT
GCCAGGAGATAGTATGACTGTAGTTCAATGGTGACTGAGCACCTGGGATGTGCTAGACA
CAAGAGTGACTTCTAAGGGTCACAGGAGAAGCTGACGTCAAAAACTTCACACAAGGGGA
CCCTGAGAGGTCACAGAAGTTCAAGATTCTGAAAGTAGTTCTGGATTCCAAGGAGCAGG
CTGGCTTCACCACTTCTGACAGGCTCTGGGAAGTAGGAGAAAGTTTGCCTCAGGTTGGA
GAGAGCAGTAGGGGAGAGGGTGGTATCCCCAAAGGGTCAGATTTCTACTCTTCTGGCAC
AAAGAAGAAGCAGAGAGGTAAAGAATAGGTCAGTATGAGCAAGGGCAACTGACCCTTTA
TGACGTAGCAAAGGAGTGGCAGCAAGTTCTGAATGTAACAAATTCTCCTTTCCTTTTTG
AAAATGTAGAACACATTAACAAATGCACT [ct] GATCAAACTGTGGTCAATCAGAAATC
GCTGCACAAAATGTCTTCCTATTAAATAAAAATCATACAGTGCTTTGCATTTGAATAGT
GTTCTATACTTTCCCATAATTCTCTCATTAGCCACCACTGGGAAATACCCTGTTATAAT
TATACAGATAAATGTGCAAATGACAGAAGAATCAATTTCTAAAAGAAGAAATACAAACT
TTTATAATGGGAGAGAGGATATATTTATTATCACTAATAAAAAAGCATATACTTCACCT
AATAAATTAATACTTTGTCACTACCAAAGTTATAATTACTATAACATTTATATATAATA
TACATTTACATTAATATTATAAATAGTAATAAATTATGAATGTTATAATTACAAATTAT
GAATTTTAAAATGTAATAACCATAATATCAATACTATATTAGTGATGGTGTTATACATT
GACACAATTTTTTTGGAAAAC

gggagagaggacctgtgacaggataaaggggctgccttatttaaacctggaaggaagaa
cgacagtataagcttccaggatattaatatcaggctaacatggacagttaagagccttt
gccaggagatagtatgactgtagttcaatggtgactgagcacctgggatgtgctagaca
caagagtgacttctaagggtcacaggagaagctgacgtcaaaaacttcacacaagggga
ccctgagaggtcacagaagttcaagattctgaaagtagttctggattccaaggagcagg
ctggcttcaccacttctgacaggctctgggaagtaggagaaagtttgcctcaggttgga
gagagcagtaggggagagggtggtatccccaaagggtcagatttctactcttctggcac
aaagaagaagcagagaggtaaagaataggtcagtatgagcaagggcaactgaccccttta
tgacgtagcaaaggagtggcagcaagttctgaatgtaacaaattctcctttcctttTTG
AAAATGTAGAACACATTAACAAATGCACT [ct] GATCAAACTGTGGTCAATCAGAAATC
GCTGCACaaaatgtcttcctattaaataaaaatcatacagtgctttgcatttgaatagt
gttctatactttcccataattctctcattagccaccactgggaaataccctgttataat
tatacagataaatgtgcaaatgacagaagaatcaatttctaaaagaagaaatacaaact
tttataatgggagagaggatatatttattatcactaataaaaaagcatatacttcacct
aataaattaatactttgtcactaccaaagttataattactataacatttatatataata
tacatttacattaatattataaatagtaataaattatgaatgttataattacaaattat
gaattttaaaatgtaataaccataatatcaatactatattagtgatggtgttatacatt
gacacaatttttttggaaaac

FIGURE 23iiii

53522
TTTTCTGTAGACTCTCCCTCCGTTTGAGCTTATCTGACATTTGCTCGCCGTGAGATCCA
GGCCTTGCATTTGTACTGGACCCTGTTCTTACACACCCTGATCCAGCCCACTTGTGTAG
TCTGGGAGTCTGGGACAACCTCCGTCCGCCCTTCTAGCCGGGTCACTGCAGGCAAGCCT
TGGTGCTCTTGCCTGCGACGTGGAAATGATGCCTGCCTGCAGCGCTGTATAGTGCAGAG
CGGGCGAGGGGCATAGGGAAGTCACTGGCACGTGGTATGTGTTGGCAGGGCTGCTTCTC
ACCCCAAACCAAGGGAGGGACAGGCAGGGAGGCTGAGAGCAGCGGCTTGCCCTGGAGCT
GTCAGGTGGGAGGCAGAGGGCGGGAGAGGCTGTGGGCTGCCCAGGTCTGATCCCTGACC
CACTTGCCACCCGTGCCCTCAGTTCTTCCCCAATGGAGAGGCCATCTGCACGGGCTCGG
ATGACGCTTCCTGCCGCTTGTTTGACCTGCGGGCAGACCAGGAGCTGATCTGCTTCTCC
CACGAGAGCATCATCTGCGGCATCACGTC[ct]GTGGCCTTCTCCCTCAGTGGCCGCCT
ACTATTCGCTGGCTACGACGACTTCAACTGCAATGTCTGGGACTCCATGAAGTCTGAGC
GTGTGGGTAAGGGCCAGCCCTGGCTGCTGCTTCCTCAGCTGGAAGGACCCTCCCCAGCC
CTCCCTCCCCATTCTGTACCCCCCATCAGCTCCCATTTCGGACTCTCTTACTGCTGTCC
CTTGTCACTGGGTGACTCCACCCCTGGAATCCAGTACCCCTTGGTTCCCAACTAGGACT
GTTTTCCCTCAGTGTTGCTCTAAGCAGCCTCTCTCCACTGCCCAATGCCATGACTGCTC
CCTGCCCTAGGAGATCTGTGGACCATGACTGTCCAGTCAGTTCTGGGTTCCTGGCATTT
CAGGGGCACCCACTGAGAGGCAAGACAGCCTCAGGGAAACATGGAATCAAGGCAGAATC
AAGGAGATCTGGAGTGGCCCG

ttttctgtagactctccctccgtttgagcttatctgacatttgctcgccgtgagatcca
ggccttgcatttgtactggaccctgttcttacacaccctgatccagcccacttgtgtag
tctgggagtctgggacaacctccgtccgcccttctagccgggtcactgcaggcaagcct
tggtgctcttgcctgcgacgtggaaatgatgcctgcctgcagcgctgtatagtgcagag
cgggcgaggggcatagggaagtcactggcacgtggtatgtgttggcagggctgcttctc
accccaaaccaagggagggacaggcagggaggctgagagcagcggcttgccctggagct
gtcaggtgggaggcagagggcgggagaggctgtgggctgcccaggtctgatccctgacc
cacttgccacccgtgccctcagttcttccccaatggagaggccatctgcacgggctcgg
atgacgcttcctgccgcttgtttgacctgcgggcagaccaggagctgatctgcttctcc
cacgaGAGCATCATCTGCGGCATCACGTC[ct]GTGGCCTTCTCCCTCAGTGGCCGCct
actattcgctggctacgacgacttcaactgcaatgtctgggactccatgaagtctgagc
gtgtgggtaagggccagccctggctgctgcttcctcagctggaaggaccctccccagcc
ctccctccccattctgtaccccccatcagctcccatttcggactctcttactgctgtcc
cttgtcactgggtgactccacccctggaatccagtaccccttggttcccaactaggact
gttttccctcagtgttgctctaagcagcctctctccactgcccaatgccatgactgctc
cctgccctaggagatctgtggaccatgactgtccagtcagttctgggttcctggcattt
caggggcacccactgagaggcaagacagcctcagggaaacatggaatcaaggcagaatc
aaggagatctggagtggcccg

FIGURE 23jjjj

53528
TGCCTCAGGTAAGAAAGACCTGGGCTTCCCTGGCTAAACGCATGAGTCCCTAGGAGGCC
AGGAAAGCCCCCAAACCCCAGCTTCGGGCCCTCCTCCCTGGCAGTGCTTCCTGGGCCCC
GGAGCCTACCCACTGAGGACTCAGTGCAGGAGTTAGGGTCTGGAGAGTATAAATGATCA
GAGTGGCTAAAAATTTCCACCACCTCCCAGTTCTCCAGGCATTTGAGTTGTGAACTCAC
CTGCTTTTTCTCCCATCTTGGACCCCCCTGGGAAATGTCCCCCTTGCCCAAGGACTGGG
CTAAAGGCCTGGGCTCATGGGATTTGGGACTCTGCAGAGGAGCAGTTCAGGGGCTGGAG
GCTCAAACCTCCAAGCAAGGACCCCTGGGCTCTCATGGCCCTGTCCCCCTTCCCAGCA
ACTAGGCTAAAGGCTGAAGGTCATGGGGACTCAACTCAGAAGGGGGGCTCGTTAGGAGC
TGAGGGGGGCCCCTCTAGGCTCTCCTGGGAGCGGGGACGGGGCAGGGCTCCTTACTGCA
GAAGGGTCTCCACCACGGCTTTCTGGTGG [ga] CCGCCTCCTCAGGGCTGAGGTTCTCC
AGCTCTTTGAGGATGGGTGGCGTGAAGTCTTCCCCATCGTCGTCCGTCTCGTCCTCGGA
GCCCCGAGTCTCCCCCAGCCCATTGGGCAGCTCAGCCAGCTCCCCTCGACCGCCGCCGC
AGGACTCCCCCTTGTCCAGGGGGCCTTCTCCAGCCAGGAGGTAGGGCCCCGGCTCACCC
AGTGCCTGGATCAGTGCCTCTTTGCTCAGCCCTGACTCGAGCAGGGCCGCCAGGAGCTC
CGTCTGCAGCTGGCTCAGTTTAGAAACCATGGCTCGGCTGCCACAGGGCCACGCGGCCC
GGGTCCACCACGCTAGCCGCCTCCCCCACCGCGTGGGTTGCGTTTGCCTGCCGGCCGGC
AGACACAAACCAAACTCCTTGCACCCACTGCCCCCCCAAAACCCCACTAGCCAAGCCCT
GTGGGCACCCCCAACCCCCAA

tgcctcaggtaagaaagacctgggcttccctggctaaacgcatgagtccctaggaggcc
aggaaagcccccaaaccccagcttcgggccctcctccctggcagtgcttcctgggcccc
ggagcctacccactgaggactcagtgcaggagttagggtctggagagtataaatgatca
gagtggctaaaaatttccaccacctcccagttctccaggcatttgagttgtgaactcac
ctgctttttctcccatcttggacccccctgggaaatgtcccccttgcccaaggactggg
ctaaaggcctgggctcatgggatttgggactctgcagaggagcagttcaggggctggag
gctcaaacctccaagcaaggacccctgggctctcatggccctgtcccccttcccagca
actaggctaaaggctgaaggtcatggggactcaactcagaaggggggctcgttaggagc
tgaggggggcccctctaggctctcctgggagcggggacggggcagggctccttactgca
gaaGGGTCTCCACCACGGCTTTCTGGTGG [ga] CCGCCTCCTCAGGGCTGAGGTTCTCC
agctctttgaggatgggtggcgtgaagtcttccccatcgtcgtccgtctcgtcctcgga
gccccgagtctcccccagcccattgggcagctcagccagctcccctcgaccgccgccgc
aggactcccccttgtccagggggccttctccagccaggaggtagggccccggctcaccc
agtgcctggatcagtgcctctttgctcagccctgactcgagcagggccgccaggagctc
cgtctgcagctggctcagtttagaaaccatggctcggctgccacagggccacgcggccc
gggtccaccacgctagccgcctcccccaccgcgtgggttgcgtttgcctgccggccggc
agacacaaaccaaactccttgcacccactgcccccccaaaaccccactagccaagccct
gtgggcacccccaaccccaa

FIGURE 23kkkk

53529
CCTCAGCCTCCCAAGTAGCTGGGACTACAGGCACGTGCCTCCACGCCTGGCTAATTTTT
GTACTTTTAGTAGAGACGGGGTTTCACCGTGTTGGCCAGGCTGGTCTTGAACTCCTGAC
CTCAAGTGATCTGCCCACCTCGGCCTTCCAAAGTGCTGGGATTACAGGCATGAGCCACC
ACGCCTGGCCCCAGATTACCTTTCTAAAATCTGAATAGATTTTAGAAATTCATATGGCC
CTAAGAGTTTCAGAGAAACACAGGCATGCACACAAATGCATGCACAACCGATACACACC
CAGACACGCACTAGGGATCTGCTCACACAAGCAGTCGTGCACACACACAGATACGTGCA
TTCACATGGGAACACACTGGCCTGCAGACACCCTCAATCACGGAAACACACTTGTCCCA
GAGACACATGCAGACTGCAATGCCTGCCAGGCACCCCTTTCCCCTGCATCCATTGACAG
CCAACCTCTATCATCATCTCCTGCTGTGTGGGGCACAGGGCGCTCACCGTGGGGGCTCT
GCAGCTGAGCCATGGTGGCCATGAAGGGG[ct]TCTGGGTCACATGGCTCTGCACAGGT
GGCATGAGCGGCTGCTGGTAGGAGGGGTGCAGCGGCTGGGAGAACTGGACGGCTGCAG
GGTGGTCAGGCTGCTGCCCATGCTGTTGATGACCGGCACACTCTGTGCCTGCGTGGAGG
CCAGGCCTGGAGTGGAAGGGGAGGGAATCAGCTGGGCCCCCCAGTTATATCCCACCCCT
GCCCAAGACCTCCCAAGGGCACCACCTCTCCTTCCCAGAGCCCGTGGTTTGGAGGAGGG
GGCAGGGTGGTCAGGAAACAGCCCTCCACTGGGACCTGCCACTAATTTAAGTGGCTCTG
GCAAGTCATTCCCCCTCTCTGAGCCTTTAGCTCTTTGTCTAGGCTAGTGGGAGAGGCAG
GCGGTGACTTGTTCAAAAGTTGTCAAACTGCGGTTCCCTGGAGCCCTGGGTTCCACAGC
AGTGCAAAGGCCATGGGGTCA

cctcagcctcccaagtagctgggactacaggcacgtgcctccacgcctggctaattttt
gtacttttagtagagacggggtttcaccgtgttggccaggctggtcttgaactcctgac
ctcaagtgatctgcccacctcggccttccaaagtgctgggattacaggcatgagccacc
acgcctggccccagattacctttctaaaatctgaatagattttagaaattcatatggcc
ctaagagtttcagagaaacacaggcatgcacacaaatgcatgcacaaccgatacacacc
cagacacgcactagggatctgctcacacaagcagtcgtgcacacacacagatacgtgca
ttcacatgggaacacactggcctgcagacaccctcaatcacggaaacacacttgtccca
gagacacatgcagactgcaatgcctgccaggcacccctttcccctgcatccattgacag
ccaacctctatcatcatctcctgctgtgtggggcacagggcgctcaccgtgggggctct
gCAGCTGAGCCATGGTGGCCATGAAGGGG[ct]TCTGGGTCACATGGCTCTGCACAGGT
GGcatgagcggctgctggtaggaggggtgcagcggctgggagaactggacggctgcag
ggtggtcaggctgctgcccatgctgttgatgaccggcacactctgtgcctgcgtggagg
ccaggcctggagtggaaggggagggaatcagctgggccccccagttatatcccacccct
gcccaagacctcccaagggcaccacctctccttcccagagcccgtggtttggaggaggg
ggcagggtggtcaggaaacagccctccactgggacctgccactaatttaagtggctctg
gcaagtcattccccctctctgagcctttagctctttgtctaggctagtgggagaggcag
gcggtgacttgttcaaaagttgtcaaactgcggttccctggagccctgggttccacagc
agtgcaaaggccatggggtca

FIGURE 231111

53530
GTGTAGATGCAGTAGCTTTTGCCTGTGGGATGGGAGGGATGGGAGATGTGTCCAGACCC
TCCTAGGAGGCCACATGAGTGTGACTGTTCTCGGCCCAAGTCTTTCTCGTTCCTCAGAG
AATTTGCGGGGCCCCTGGGCACACAAGCTGAGATCCACCCAGCCCTGGTCCCTTGGCAA
GAACTGAGGGACAGGACCTGGTTCTGGGGAAAATGCAGGGGAATGTTTCTCCCTTCCAC
AGCCCCCTTGCGAGTTAGGAGGCCGGCTCCCACCCCAGAAGGTGGCCAGGTTTTCATGC
CTTCCTAGAGAAAGCTGGGGCTCGTGGCCTCCACCACAGGAAGACGCAGACCCTCAGAA
ACAAGTCTGTGAAGTCACAACCAGCCCCAGTTTACAGATGTGAAACTGAAGCTCCAAAA
AGTCAGGAGGTCACTGAGTGGGGAGGTGATGGAGTGGGAACAGCCCCCAGATCTGGCTG
AGGCCGAAGCCCTGGAGAGATCCCCGCAAGGCTCCCTTAGATGCCTGACATTCTGTTCT
TCCTGAAGCCTCACTCCCTTCTCTCCTGG[ct]GCAGACACGTCCCCATCAGAAGGCAC
CAACCTCAACGCGCCCAACAGCCTGGGTGTCAGCGCCCTGTGTGCCATCTGCGGGGACC
GGGCCACGGGCAAACACTACGGTGCCTCGAGCTGTGACGGCTGCAAGGGCTTCTTCCGG
AGGAGCGTGCGGAAGAACCACATGTACTCCTGCAGGTGAGGAGCCTCAATTTCTTCAGC
TGGGAAATGGGCACACTTGGGCTCATGGCCCCAAGGTCTGTCTTCTCCCTGAGTGGGTA
GGTCCCAGAGACAGCTGCCCTTCAGGGCCTTCAAGGCTCTTCTGGTTTTGTAAAAGACT
TTGTGAATCCAAGAAGAGCATCTATTCTAGGAACCACATTTACTGATCATCAAGCTACT
GGCTGCCGTTTATTGAGCTCTTATCATATGCCAGGCACAATACTAAGTCTTTGTGTGTA
TTTACCCATCCCCTTGAGCCC

gtgtagatgcagtagcttttgcctgtgggatgggagggatgggagatgtgtccagaccc
tcctaggaggccacatgagtgtgactgttctcggcccaagtctttctcgttcctcagag
aatttgcggggcccctgggcacacaagctgagatccacccagccctggtcccttggcaa
gaactgagggacaggacctggttctggggaaaatgcaggggaatgtttctcccttccac
agcccccttgcgagttaggaggccggctcccaccccagaaggtggccaggttttcatgc
cttcctagagaaagctggggctcgtggcctccaccacaggaagacgcagaccctcagaa
acaagtctgtgaagtcacaaccagccccagtttacagatgtgaaactgaagctccaaaa
agtcaggaggtcactgagtggggaggtgatggagtgggaacagcccccagatctggctg
aggccgaagccctggagagatccccgcaaggctcccttagatgcctgacattctgttct
tcCTGAAGCCTCACTCCCTTCTCTCCTGG[ct]GCAGACACGTCCCCATCAGAAGGCAC
Caacctcaacgcgcccaacagcctgggtgtcagcgccctgtgtgccatctgcggggacc
gggccacgggcaaacactacggtgcctcgagctgtgacggctgcaagggcttcttccgg
aggagcgtgcggaagaaccacatgtactcctgcaggtgaggagcctcaatttcttcagc
tgggaaatgggcacacttgggctcatggccccaaggtctgtcttctccctgagtgggta
ggtcccagagacagctgcccttcagggccttcaaggctcttctggttttgtaaaagact
ttgtgaatccaagaagagcatctattctaggaaccacatttactgatcatcaagctact
ggctgccgtttattgagctcttatcatatgccaggcacaatactaagtctttgtgtgta
tttacccatccccttgagccc

FIGURE 23mmmmm

53540
ATACACCAAATTTGTTTACTTTGAATAGCTTTCTTTGGACAGAGGAATTTTGAGTACTT
AATATTTTTTGCATATTTTTCATACTTTCCATCATGAACATGTATGGCTTTTACATTTA
GGAAGAAATAATGCTATTTTTTAAAGGAGGAAAAAAGAGAAAAGAGTTGGTGCGAATAA
TTGAAGTAATCTATTATGCAGTGTGTGAGTAATGAATTGATAGATAGGATCATCTGTAG
ATTTCAAGGAGCTATAATTTCCCCTGTAACATGTTTTTCAACATTTCTCTCCCCTTTTA
TTATAAAAAACACAAACTCTGATCTACACTCCAACAAAGTCTGCTTTTATCACAAGGAT
ACTTTAAACATTTGATCATTGTGCAGAATATTTATTCTAAATTACTGAGACCTTATTCA
CTAATCATAGTTTTCACAGGCTTTATTCCAACCATATTGATATGTTAGTTCGAGACTAC
GGATTTAATACCTGGATTTCTCCTCTGTGTCTTGAAGGGAACGTTGCCAGCTGCCTTGT
ACCAGCATTACAAATAATCCAGCCACAAA [ga] TAAATGCTTTTCATTTCTGCTGTCTG
TCAGAACACAGAATGGGGTAGGGTGAGGGGGGCAGGCAAGGATTTTTAAACATGTCAG
GCTAAATTAATTAGATTTGACTAGATAAATATCATAAGTAGAAGGAAAAAGCTAGTGTT
ATCACTTTTATTCTGATTATATTTTCAGCTTAATTTTAAATAGTGGGTTATATTATTTC
CCCAGATTTTTTGGAGGCAAAAAAGGACACAAAAGATGTGTTCCACCATTAAGCTTTTT
CATTAATGTAGGGACACTTCTGTTTAATAATTAGAAGGCTCATTTCCAGACTGGAAATT
AAAATGTCCACAATCAACATTTAAAATACCCACTGTAGATGATATGCTACATATGGTTA
GCCTGAATGGCACCTTATCCATCATGCCACCCCCCTCACTATCAGTCTGGCTTTCAATT
AATAGTCCTTCACTTCCAAGC

atacaccaaatttgtttactttgaatagctttctttggacagaggaattttgagtactt
aatattttttgcatatttttcatactttccatcatgaacatgtatggcttttacattta
ggaagaaataatgctattttttaaaggaggaaaaaagagaaaagagttggtgcgaataa
ttgaagtaatctattatgcagtgtgtgagtaatgaattgatagataggatcatctgtag
atttcaaggagctataatttcccctgtaacatgtttttcaacatttctctcccctttta
ttataaaaaacacaaactctgatctacactccaacaaagtctgcttttatcacaaggat
actttaaacatttgatcattgtgcagaatatttattctaaattactgagaccttattca
ctaatcatagttttcacaggctttattccaaccatattgatatgttagttcgagactac
ggatttaatacctggatttctcctctgtgtcttgaagggaacgtTGCCAGCTGCCTTGT
ACCAGCATTACAAATAATCCAGCCACAAA [ga] TAAATGCTTTTCATTTCTGCTGTCTG
TCAGAACACAGAATGGGGgtagggtgaggggggcaggcaaggatttttaaacatgtcag
gctaaattaattagatttgactagataaatatcataagtagaaggaaaaagctagtgtt
atcacttttattctgattatattttcagcttaattttaaatagtgggttatattatttc
cccagattttttggaggcaaaaaaggacacaaaagatgtgttccaccattaagcttttt
cattaatgtagggacacttctgtttaataattagaaggctcatttccagactggaaatt
aaaatgtccacaatcaacatttaaaatacccactgtagatgatatgctacatatggtta
gcctgaatggcaccttatccatcatgccacccccctcactatcagtctggctttcaatt
aatagtccttcacttccaagc

FIGURE 23nnnn

53544
TAGGAATTGTGCATCAGGAAAGTGAAGAGGATTGCTAGACATTTAGTCCTGTTATAAGA
GCACTAAAGATTTGGCAGTCACCAGGTATGGAGTCTCAGGAGGAGCTTACCGATGGATG
GGGCATAGCCATTATATTTGCCCGAGTCCAGGGCATCTTTCATTGCCTGGGTAACTTCA
GGGTCTGTAGGCAGGTTTCCAAACACAGTAGGGTCCCCTTTTTATGGGAGGAAAACACA
AAAGGAGCCAAGAGGTTATTCTCCCATGTTCAGTACTCAGACTCACCCCCAACTGCCAT
CTTCTCCAACCAGCCTGTGAACATGAGAGTAGAGGAGGACAATGACAGCCCCTCAGTAG
TGTCCCCAACTCACCAATGGACAGGGAAATCATGGTTTTGTTTGGATTTGGTTTCACCT
TCATGTTGTCCACAATGGCTCGGATGGGGTTGAAAGTTTTCTTGGCCATGTCTGAGGGC
CTCACAGACCACCTGGCCTTTCTGCCTTTCATTTTTCCCGGCACAGAGCTTCTCCCACC
AACGTTGACATGCACGTCCAGAATTGAGG[ga]GAGGTTGCCTTTGCTGCTCATCTGAA
TCATGTATGGGTCCATCACTAGCGAAGCCTGCGAGGGGAAAGAAGTTCCCTGTGATGTT
GATAACATAGCGCTGGGGGACAGAGGAGCTACATTTGGACCTAAACATTGGGTGACTTC
ACTAAAAGTGTCTTTCCAAACTCTCTCTTTATTTTTTTTCTACTTTCTGTTGTAAAGT
AGCTTTACTATGAATGGGGGAGTTTTAAGAGTTTTTACTGAGATGGAAAATAAAGCAAG
AACCCATTCTACTTAAGTAGGATTTGCTACACGCATCTGCAATTCCTGTCAAAGCTTAA
CCATGCTCTATGTGAAACCAAGAAGGAATAAGATGAAAATTGTTCATCAGTCAAAGCAT
AGGTTCTCCTTCCTTTCCATGCGAGCCTATCCAAGAAAATCTACCTAATGCTTCTTGTC
ATCTGCAGAGGACCAGGAAGA

taggaattgtgcatcaggaaagtgaagaggattgctagacatttagtcctgttataaga
gcactaaagatttggcagtcaccaggtatggagtctcaggaggagcttaccgatggatg
gggcatagccattatatttgcccgagtccagggcatctttcattgcctgggtaacttca
gggtctgtaggcaggtttccaaacacagtagggtcccctttttatgggaggaaaacaca
aaaggagccaagaggttattctcccatgttcagtactcagactcacccccaactgccat
cttctccaaccagcctgtgaacatgagagtagaggaggacaatgacagcccctcagtag
tgtccccaactcaccaatggacagggaaatcatggttttgtttggatttggtttcacct
tcatgttgtccacaatggctcggatggggttgaaagttttcttggccatgtctgagggc
ctcacagaccacctggcctttctgcctttcatttttcccggcacagagcttCTCCCACC
AACGTTGACATGCACGTCCAGAATTGAGG[ga]GAGGTTGCCTTTGCTGCTCATCTGAA
TCATGTATGGGtccatcactagcgaagcctgcgaggggaaagaagttccctgtgatgtt
gataacatagcgctgggggacagaggagctacatttggacctaaacattgggtgacttc
actaaaagtgtctttccaaactctctctttattttttttctactttctgttgtaaagt
agctttactatgaatgggggagttttaagagtttttactgagatggaaaataaagcaag
aacccattctacttaagtaggatttgctacacgcatctgcaattcctgtcaaagcttaa
ccatgctctatgtgaaaccaagaaggaataagatgaaaattgttcatcagtcaaagcat
aggttctccttcctttccatgcgagcctatccaagaaaatctacctaatgcttcttgtc
atctgcagaggaccaggaaga

FIGURE 230000

53550
CCAACTAGAATACAGCTTCCTGAGAGGCAGGATCTTGACTGACTTGTTCATTCCTAATT
TCTCAGCATCTAGAAGAAGGTATGGCACATAGTAGGTGCTTGTTAGATACTTGCTAATA
AATGGAAATAAACATATCCCTAGTTCCTATTCCAGCTTTTTCCCTGCTGTTTTGTCCTC
CATTCTTCCAGCAGACAACAGGACTAGTTCCCTGACCCCCTGCAGGAAGCTAACAATAC
CCTAGCCTACTTCTAAGCAAAACGTCGCAGCTTCAAAGACTTTCCATGGAGGGCGATGG
GCTGAGGACAATCTTGTTCTTCACGTAAAACACAGGCCCACAATCTCAAATTTATAATT
TAAAAATATATATACTTACAATGTCTCTAAAGGCACTTATTTTTCTTAAAAATCATGTA
TTTGTAAGCTGAACTATCATTTTAACACAAAAGCTATCATTCTTGCTCAATGGAGTCAG
GCTGCTCTTGGAGTTTCTGTCCTGGGAGGAAAAAGGGCAGGGTGTAGGTACCTGATGGT
TTTCCACA[ct]GTCGAAGCCATCCAGAGGCTTTGTGCCATTGGTGTGTCCCCTGGCCA
GCTTCACGAGTGTTGGCAGCCAGTCAGAGATGTGGATGAGCTCCCGGTTCTTCACGCCC
TTCTGCTTCAGCAAGGGGCTTGCCACAAAGCCCACCCCTCGGACGCCTCCTTCCCACAG
GCTCCATTTTCTTCCTCGAAGGGGCCAGTTATTACCCCCTGCCAAAGTCTGCCCTCCGT
TATCTGAAACACAGTAAGGTCTTGGCATGAGGATGATGTTAACTCTTAAATACATTTAA
GAACAGAGACTGTATGTACATTGTTACTAAATGGTGCTTAAATAATAAAAAAAAAGAAA
ATTCCTTGCCTTTTCCCACCCTAAATTCCCTTTTCCCATTGACATAGCCTTTCATTATT
CAGACATAAGTAAGGCCCAGTGTGATACATATCTACCTTTAAATCCTCCATGGAGAGAG
CCACTGGAAAACAAGGCAGTC

ccaactagaatacagcttcctgagaggcaggatcttgactgacttgttcattcctaatt
tctcagcatctagaagaaggtatggcacatagtaggtgcttgttagatacttgctaata
aatggaaataaacatatccctagttcctattccagctttttccctgctgttttgtcctc
cattcttccagcagacaacaggactagttccctgaccccctgcaggaagctaacaatac
cctagcctacttctaagcaaaacgtcgcagcttcaaagactttccatggagggcgatgg
gctgaggacaatcttgttcttcacgtaaaacacaggcccacaatctcaaatttataatt
taaaaatatatatacttacaatgtctctaaaggcacttatttttcttaaaaatcatgta
tttgtaagctgaactatcattttaacacaaaagctatcattcttgctcaatggagtcag
gctgctcttggagtttctgtcctgggaggaaaaaggGCAGGGTGTAGGTACCTGATGGT
TTTCCACA[ct]GTCGAAGCCATCCAGAGGCTTTGTGCCATTGgtgtgtcccctggcca
gcttcacgagtgttggcagccagtcagagatgtggatgagctcccggttcttcacgccc
ttctgcttcagcaaggggcttgccacaaagcccacccctcggacgcctccttcccacag
gctccattttcttcctcgaaggggccagttattaccccctgccaaagtctgccctccgt
tatctgaaacacagtaaggtcttggcatgaggatgatgttaactcttaaatacatttaa
gaacagagactgtatgtacattgttactaaatggtgcttaaataataaaaaaaaagaaa
attccttgccttttcccaccctaaattccctttttcccattgacatagcctttcattatt
cagacataagtaaggcccagtgtgatacatatctacctttaaatcctccatggagagag
ccactggaaaacaaggcagtc

FIGURE 23pppp

53551
CTGTGGGGAGCGTGGCTTTGCTACTCAATGGCAACTGGATTTCAAGAGTTTCAGGAAGG
GTGGGGGAGCAAGATATCAAAGGCTCAAGCTCACTCCCCTTCGTCCAGACAGACTTTTC
ATTTTTTGTTTGATGAAGATTAGGAAGAAAAGAGTGAGGATTAGGCCTAATTTACTGCC
TCTGTCAAAAGCCAGCGCAGAGTAGAAGGGAAGGGAGTAAGTGGATTATGAAAAGAAAA
CAAACGGAGGGAAAGGGGGCCGAGGATGAACTGCATTCAGTGATATTTATTTATCTGAT
TGCAAAAGGAAAAGAAGGGATCTGTTCTAATGGTTCACCTTCTTATGAACCCTGGAGCT
CCCAAAACCCTGGCGAAGTCCTTCTGACACTGCTGTGAGGTAGATCGGAGCCATTCCAT
GGCTAAAGTGAGAGAGGCCACTGCTTGAGAGCAGTAATAAGGGAACCAGAGATAAAACC
CCAAATCTTGGTCTTTTCTACCCTGCTGCTCTCAGCCTGGGCCACAGAGCCTGGAGAAC
ACTAAGGTCTCATCAGGGTTTGGGTGGCA[ga]AAGGAATGGAACCAGGGGAGCTCTCT
TTGCCCTAAGCACTCACTGACTGCACAGGCAAGCCGGGTGATGGGTGCCCCTACCAAAG
CCAGCCTGCTGCTCCACGGCACCTGGACACTACCACTGAGGGAGGAGTGAAGTTCAAGG
CTGGGGTTTAGAAAACATCTCTCAGACAGAGAGCAAGAGGATGGTGAAAACCCACTTGG
TAAGGATCCCTCCTTGGGTCACATGGCCCAGTCGTCAGGTTCTGGAGGGTAGAGTGTCA
CAGCCGGGGAATCCCATGGGACTCATTCTGAACAGAGGCCAGAGGTTTTCCACAGGTTC
TGATCAACAGAGTTGTTGCTTCTTGTCCTTCAGGCCTAAGAAACTCCCCAAGAAGCCCT
GGGAAAAAAAGTGGAGATAATAGACCCTGGGGTGAAAGGAGCAACAGGTGCACTGAGGG
GAATGACAGAGATCAGAGACC

ctgtggggagcgtggctttgctactcaatggcaactggatttcaagagtttcaggaagg
gtgggggagcaagatatcaaaggctcaagctcactccccttcgtccagacagacttttc
atttttttgtttgatgaagattaggaagaaaagagtgaggattaggcctaatttactgcc
tctgtcaaaagccagcgcagagtagaagggaagggagtaagtggattatgaaaagaaaa
caaacggagggaaagggggccgaggatgaactgcattcagtgatatttatttatctgat
tgcaaaaggaaaagaagggatctgttctaatggttcaccttcttatgaaccctggagct
cccaaaaccctggcgaagtccttctgacactgctgtgaggtagatcggagccattccat
ggctaaagtgagagaggccactgcttgagagcagtaataagggaaccagagataaaacc
ccaaatcttggtcttttctaccctgctgctctcagcctgggccacagagcctggagaAC
ACTAAGGTCTCATCAGGGTTTGGGTGGCA[ga]AAGGAATGGAACCAGGGGAGCTCTCT
TTGCCctaagcactcactgactgcacaggcaagccgggtgatgggtgcccctaccaaag
ccagcctgctgctccacggcacctggacactaccactgagggaggagtgaagttcaagg
ctggggtttagaaaacatctctcagacagagagcaagaggatggtgaaaacccacttgg
taaggatccctccttgggtcacatggcccagtcgtcaggttctggagggtagagtgtca
cagccggggaatcccatgggactcattctgaacagaggccagaggttttccacaggttc
tgatcaacagagttgttgcttcttgtccttcaggcctaagaaactccccaagaagccct
gggaaaaaaagtggagataatagaccctggggtgaaaggagcaacaggtgcactgaggg
gaatgacagagatcagagacc

FIGURE 23qqqq

53553
CTTTAGAAACGGCTCTAGGTTGAGACCGCCGGCATGGATCTCCACCTCTACTGCAGACA
CACACTGGAAGGCTTCGGACCAGTCGGGCTGAGGTTCGGAGAAGTTGCAGACGCAGCGG
AAATCTTCATCGTCCAGCTCACAAGGTTCTGGCGTGGTCGCAGAGACGTGCACCAGCGG
CAGCAGCAGCAGCAACAAGCAGGACGCGCGCTCCTGGGGAGAGAGCAGAGGTCTAGGAG
GCCCCATCCAACCCCTGTGGCTCCCGAGTGGCACGCGTTCGACCCCAAGACCCTACACT
CACCATGGTCGATAAGTCTTCCGAACCTCTGAGCTCCGGACAGGCTCTGGAAGTGCTTT
ACGTTCTTTCCTACACAGCGGCACCCGCCGGCTTCCAGGCTTCACACTTGTGAACTCTT
CGGCTGCCTCTGACAGTTTATGTAATCCTGGGATGTCATTCAGTTCCCTCCTCTGTGAA
CCCTGATCACCTCCCCACCTCTCTTCCTCCGAGCCAGCCCCCTTCCTTTCCTGGAAATA
TTGCAATGAAGGATGTTTCAGGGAGGGGG[ag]CCGTAACAGGAAGGATTCTGCAGGGC
ATCTAGGGTTCTGTGTCTCCTGGCAGTGTCCTGATGACTCAGGCGCCCCAGGCGGTGAA
TGCCCTGTTGACTCGGGAGCCTAAGCCTTCTCTGGTGGGTGTGGGAAAAGGATGATCCT
CAGTGCCTTAGGCCAGTACCATACTCTGCACTATCCAACCCCCCAATCCCCCTACCTTA
TATCCCAGAGAATCTACTTGATTCATTTCTTTGACTTCTTCCTTGTCTTGGTTTATGTT
GATCTCCTGCCACCAAATCCAAGTCCCTGAATATCCTCAGATATTTAACTGCATGTTTT
GTGGAAGAGATTGTGAACCTCATCTGTTGGCACCAAGGGGGGTAGAATTAGGTTCAAGA
AAAGGAAGTTGGTCTAAAGAAAAATTCCCCCTTCCTTTTTTTTTCCTTGCTCCTTTGAT
TAAGTAATAACTTTCTTTCTT

ctttagaaacggctctaggttgagaccgccggcatggatctccacctctactgcagaca
cacactggaaggcttcggaccagtcgggctgaggttcggagaagttgcagacgcagcgg
aaatcttcatcgtccagctcacaaggttctggcgtggtcgcagagacgtgcaccagcgg
cagcagcagcagcaacaagcaggacgcgcgctcctggggagagagcagaggtctaggag
gccccatccaacccctgtggctcccgagtggcacgcgttcgaccccaagaccctacact
caccatggtcgataagtcttccgaacctctgagctccggacaggctctggaagtgcttt
acgttctttcctacacagcggcacccgccggcttccaggcttcacacttgtgaactctt
cggctgcctctgacagtttatgtaatcctgggatgtcattcagttccctcctctgtgaa
ccctgatcacctccccacctctcttcctccgagccagcccccttcctttcctggaAATA
TTGCAATGAAGGATGTTTCAGGGAGGGGG[ag]CCGTAACAGGAAGGATTCTGCAGGGC
ATCTAGGgttctgtgtctcctggcagtgtcctgatgactcaggcgccccaggcggtgaa
tgccctgttgactcgggagcctaagccttctctggtgggtgtgggaaaaggatgatcct
cagtgccttaggccagtaccatactctgcactatccaacccccccaatccccctacctta
tatcccagagaatctacttgattcatttctttgacttcttccttgtcttggtttatgtt
gatctcctgccaccaaatccaagtccctgaatatcctcagatatttaactgcatgtttt
gtggaagagattgtgaacctcatctgttggcaccaaggggggtagaattaggttcaaga
aaaggaagttggtctaaagaaaaattccccct tcctttttttttccttgctcctttgat
taagtaataactttctttctt

FIGURE 23rrrr

53557
GCAGGGCTCAGCCTGCCTCCCTGCTGCTGAGGCCCCTACCAAATTGGAACCCGAGTAGC
ACCAGGGAAGCAGGGCCTGCAGGGGATGCCATTCTCACCCCTGCCTGCAAAACGCTGCA
GTGCCCGAGTCTGCTGTGGGCTGGTGGGGGAAGGGCATCGCTAGGTTGGTGGCTGCCCC
CACCCCAGCACACTCCCCCCATTCTCTTTAGATTGTCTCACAGGGGGACCCACTTGGTT
CTCATTCTGAACTTTCAGTGAATGGATTCTGCTCCCTGCCTTGCGTGTGTACCCTTGGG
TGGCCTTTGCCCGTATCTTAGTCTCAGTTTCCTGAGTTTGGGCAGGAAGGAGAGGAGGG
GTTCTGACTGATGAGTTACCTCTTCTCCCTCTCCCCACCTCGCAGGGGGCTCCTGAGAG
TGTGATCGAGCGCTGTAGCTCAGTCCGCGTGGGGAGCCGCACAGCACCCCTGACCCCCA
CCTCCAGGGAGCAGATCCTGGCAAAGATCCGGGATTGGGGCTCAGGCTCAGACACGCTG
CGCTGCCTGGCACTGGCCACCCGGGACGC [cg] CCCCCAAGGAAGGAGGACATGGAGCT
GGACGACTGCAGCAAGTTTGTGCAGTACGAGGTGGGTGCAGGAGCCGATTCTCCCTGCA
GTACGAGGTGGGTGCAGGAGCCAAGTCTCCCTGCAGCAGCTGAGCAGGTGGTAGGTCAG
GGATGGGCTCAGGCCCCGCTTGAATCTGCCCCCTCCCTACAGACGGACCTGACCTTCGT
GGGCTGCGTAGGCATGCTGGACCCGCCGCGACCTGAGGTGGCTGCCTGCATCACACGCT
GCTACCAGGCGGGCATCCGCGTGGTCATGATCACGGGGGATAACAAAGGCACTGCCGTG
GCCATCTGCCGCAGGCTTGGCATCTTTGGGGACACGGAAGACGTGGCGGGCAAGGCCTA
CACGGGCCGCGAGTTTGATGACCTCAGCCCCGAGCAGCAGCGCCAGGCCTGCCGCACCG
CCCGCTGCTTCGCCCGCGTGG

gcagggctcagcctgcctccctgctgctgaggcccctaccaaattggaacccgagtagc
accagggaagcagggcctgcaggggatgccattctcacccctgcctgcaaaacgctgca
gtgcccgagtctgctgtgggctggtgggggaagggcatcgctaggttggtggctgcccc
caccccagcacactccccccattctctttagattgtctcacagggggacccacttggtt
ctcattctgaactttcagtgaatggattctgctccctgccttgcgtgtgtacccttggg
tggcctttgcccgtatcttagtctcagtttcctgagtttgggcaggaaggagaggaggg
gttctgactgatgagttacctcttctccctctccccacctcgcagggggctcctgagag
tgtgatcgagcgctgtagctcagtccgcgtggggagccgcacagcacccctgaccccca
cctccagggagcagatcctggcaaagatccgggattggggctcaggctcagacacgctg
cGCTGCCTGGCACTGGCCACCCGGGACGC [cg] CCCCCAAGGAAGGAGGACATGGAGCT
GGacgactgcagcaagtttgtgcagtacgaggtgggtgcaggagccgattctccctgca
gtacgaggtgggtgcaggagccaagtctccctgcagcagctgagcaggtggtaggtcag
ggatgggctcaggccccgcttgaatctgccccctccctacagacggacctgaccttcgt
gggctgcgtaggcatgctggacccgccgcgacctgaggtggctgcctgcatcacacgct
gctaccaggcgggcatccgcgtggtcatgatcacgggggataacaaaggcactgccgtg
gccatctgccgcaggcttggcatctttggggacacggaagacgtggcgggcaaggccta
cacgggccgcgagtttgatgacctcagccccgagcagcagcgccaggcctgccgcaccg
cccgctgcttcgcccgcgtgg

FIGURE 23ssss

54762
AGCAGAGAAGACAAATAATAGATACTGCGAAGATAGGATGATTGAAGAATGCAGTGATA
TAAATTTGGGGGAAGAGGAGGGAGGCAGAGCAAAGAAATTCAAGGCCTTGGCCAGACGT
AATGTCTCACACCTTGTAATCCCAGCAGTTTGGGAGGCTGAGGCAGGCTGATAGCTTGT
GTCCAGGAGTTCGAGACCAGCCTGGGCAATCCAGCAAAACCCTGTGTCTACAAAAAAAT
ACAAAAATTAGCCAGGCATGGTGGCATGCGCCTGTGGTCCCAGCTACTTGGGAGGCTGA
GGTGGGAGAATCGCCGGGACGTCGAGATTGCAGTGAGCTGAGATCGTGCCACTGCACTC
CTGCCTGGGTGACAGAGCAAGAACGTCTCAAAGAAAAAACAACAACAACAACAACAACA
ACAACAACAAAAACACAAGGCCTGTGGTTGGGGGAAGGTTGTAACTCTAAAAAAGACCC
ATGTGGCTACAGCGAGGGACACTGGGTGTAGGTAGAGATAAGAAGAGTGATACTCAGTT
CTCACATCACGGCGGACTGAATACAGGCC[cg]GGGGAGTGAGAGACCATCCACCCCTG
TGATCTGGGGCAAGTCACCAGCCCTTTCAGAGAAGCTTCCGTCTTCTCTGCAAAATGGG
ACAATACCTTGCTTCACAAGCTTGCAAGGATCAAAAGAACTGGTAGTGGCCGGGCGCG
GTGGTTCACGCCCGTAATCCCAGCACTTTGGGAGGTCGAGGCAGGTGGATCACTTACTT
GAGGTCACGGGTTCGAGACCAGCCTGGGCAAAATGGTGAAACCCCGTGTCTGCTAAAAA
TACAAACATTAGCCTGGCGTGGTGGCAGGTGCCAGTGATCCCAGCTACTCGGGAGGCAG
AGGCAGGAGGATCGCTTGAACCCAAGGGGTGGAGGTTGCAGTGAGCTGAGATCGCGCGC
TGCACTCCAGTCTGGGCAACAGATCAAGACTGTCTCAGAAAAAACAAACAAAAAAGAAC
TGGTAGAGGAAGCGCTTTGCA

agcagagaagacaaataatagatactgcgaagataggatgattgaagaatgcagtgata
taaatttgggggaagaggagggaggcagagcaaagaaattcaaggccttggccagacgt
aatgtctcacaccttgtaatcccagcagtttgggaggctgaggcaggctgatagcttgt
gtccaggagttcgagaccagcctgggcaatccagcaaaaccctgtgtctacaaaaaaat
acaaaaattagccaggcatggtggcatgcgcctgtggtcccagctacttgggaggctga
ggtgggagaatcgccgggacgtcgagattgcagtgagctgagatcgtgccactgcactc
ctgcctgggtgacagagcaagaacgtctcaaagaaaaaacaacaacaacaacaacaaca
acaacaacaaaaacacaaggcctgtggttgggggaaggttgtaactctaaaaaagaccc
atgtggctacagcgagggacactgggtgtaggtagagataagaagagtgatactcagtt
cTCACATCACGGCGGACTGAATACAGGCC[cg]GGGGAGTGAGAGACCATCCACCCCTG
TGatctggggcaagtcaccagccctttcagagaagcttccgtcttctctgcaaaatggg
acaataccttgcttcacaagcttgcaaggatcaaaagaactggtagtggccgggcgcg
gtggttcacgcccgtaatcccagcactttgggaggtcgaggcaggtggatcacttactt
gaggtcacgggttcgagaccagcctgggcaaaatggtgaaaccccgtgtctgctaaaaa
tacaaacattagcctggcgtggtggcaggtgccagtgatcccagctactcgggaggcag
aggcaggaggatcgcttgaacccaaggggtggaggttgcagtgagctgagatcgcgcgc
tgcactccagtctgggcaacagatcaagactgtctcagaaaaaacaaacaaaaaagaac
tggtagaggaagcgctttgca

FIGURE 23tttt

54777
AAAAAAAAAGTGGCTGGAACTGCCATCACTATCCTAGAGATGGAAGGTTAGGCCAATGC
TACAGCAAGGTAGCTGTGGTCAGACACTAAGAATGCTCCTTCTATCTGGCTGCCAGCCA
ATGGATCTCCATTCTGGACCAGCCCACGAGAAGCAAACCTCAAAGGAAACTAATCTGAG
GTCTTAGCTCAATCTGTGGGGAACGGCATTAAAGCCTCTCCCTCTGAGTGACCTCTGCT
AGCTTCTCTACCTCCTGCTTCCTCATCTGCTTCTGCTACACACCCGCACACTGAAAACC
CTGTATATTGTATGAGTCCTCCCTGAACCCCACATCAGTCCTGAGGTGCAATTCTGCCT
AGTCATCTTTCCTCTTCCCTCAACAGCAGCTTACTTTATGTTCTTCAAGCTTCACTGAG
GCCTCTTTTGCAAATCCTCCCAGATCTCCTCAGCTGGGATGGGGCCCCTCTAGGCTTCC
TGAGCCCCATGCTTCCTCCCTTCATGGCATCTGTCATAATGCAGTGGGATTGCCATGTA
ACTCCCTTGACTGTCTCCCCAACACAGAG[ct]TGTACACTTCACATCTGGGCAGGGTC
ACCATGACTGTGTCCACCATTGCCAGCTTGGAACCTGGCATACTGGCATCAGTAAATGT
TTGCTGAAAGAATAAATGATAACAAGCTGTCCTGCCCACCGTGACCTTTGGGAGAATGG
GCATATGCTTTTGATTACCTGCAGGGCCATCAAGGTGTTGGCCAGGGCTTGACCATAGG
TGTCATGGCAGTGGACAGCCAGGGCAGCCAGAGGCACTTCCTGCATGACAGCAGATAGC
ATGTCTTTCATGATCCCTGGGGTGCCCACACCAATGGTGTCCCCCAGGGAGATCTCGTA
GCAGCCCATTGAGTAGAACTTCTTGGTGACCTAAGGAAGCAAGCAGGCACTTGGAGGAT
ACAGAATCCACCAGCCAGGGGATCCATGCACTCAGAAGAGGGGGCCTTTGCCTGGGCAG
AACACTTCTGGGTATGACGCA

aaaaaaaaagtggctggaactgccatcactatcctagagatggaaggttaggccaatgc
tacagcaaggtagctgtggtcagacactaagaatgctccttctatctggctgccagcca
atggatctccattctggaccagcccacgagaagcaaacctcaaaggaaactaatctgag
gtcttagctcaatctgtggggaacggcattaaagcctctccctctgagtgacctctgct
agcttctctacctcctgcttcctcatctgcttctgctacacacccgcacactgaaaacc
ctgtatattgtatgagtcctccctgaaccccacatcagtcctgaggtgcaattctgcct
agtcatctttcctcttccctcaacagcagcttactttatgttcttcaagcttcactgag
gcctcttttgcaaatcctcccagatctcctcagctgggatggggcccctctaggcttcc
tgagccccatgcttcctcccttcatggcatctgtcataatgcagtgggattgccatGTA
ACTCCCTTGACTGTCTCCCCAACACAGAG[ct]TGTACACTTCACATCTGGGCAGGGTC
ACCATGactgtgtccaccattgccagcttggaacctggcatactggcatcagtaaatgt
ttgctgaaagaataaatgataacaagctgtcctgcccaccgtgacctttgggagaatgg
gcatatgcttttgattacctgcagggccatcaaggtgttggccagggcttgaccatagg
tgtcatggcagtggacagccagggcagccagaggcacttcctgcatgacagcagatagc
atgtctttcatgatccctggggtgcccacaccaatggtgtcccccagggagatctcgta
gcagcccattgagtagaacttcttggtgacctaaggaagcaagcaggcacttggaggat
acagaatccaccagccaggggatccatgcactcagaagagggggcctttgcctgggcag
aacacttctgggtatgacgca

FIGURE 23uuuu

54874
ATTGGCCTTGTTCCCCAGGGTGGAGCTGTCACAAAATAGAGTGGGAACTGTCTGGCTTT
CAGCCCAAGAGAATCTGCATGGCAAGTTGCATTAACAACCAGGCATTTCCGGCAGTTCC
CAACATTTCTGGGAATTTTCTCATCCAAACGACTGAAAGCCCACTCCATTCTCTTGCTT
CTTACTCATGCTTTCTTTGTATAATGGTAATTATGTTTTAAAAAATCCTGGGCTATGTT
GTTTCATGGAACAATTTAGAACTTATTGGTCAAACTCTGAAGCAAAGGTATATAAAAGG
TAGTTAGAGATGTTTAGGGAATATTCAAAGCACATTTTTGGGTCACTCATAATTGATCT
TTATATTCATATATGTATATATATATATATAACATAATGTACCCATCTTAACATATCAA
AGCTAAACCAGTATTAAAAACAACTGACTATGGTCTATTGATACAATATATGATGCCCA
AGTACACTCTTCATTGCTACTGCATATCTAAAATCATTTATTTATTTATCCATCCATCA
AGAGTGTATTGAGAGCCTGACAACATACC[ag]GCATCAAGCCCTGGAGGTCTTTTTAA
GGCTGAGCCAATATAGCTATGGATAACATTCTAAAACTGATAGCATATTTTCATGTTTT
ATAGTCTTTCCACAGACTAGTTCAAAATGAACACTGCCTGAGAGGGGCTTTAAGATGAC
TGACTAGAGGTACTGGACACCTGTTTCCCCAGCAAAGAAGAGCCAAAATAGCAAGTAGA
TAATCATACTTTGAATAGACATCTAAGAGAGAATGCTGGAATTCAGCAGAGAAGTGACA
GAAAACACCTGAGATACTGAAGGAGAGGGAGGCAAGGTAGACAGCCTGGCTGGAATCAG
CTGGGAGCCCAGAGAGGGTCCCTAGTGAGAGGAAAGGGTAAGTGAGAGATTCCCAGTGG
TACATGTTCCCATGTTGACTGCTGAAATCCTAGTCATAAGAGTCTCTCAAACCCCAAGG
ACCCTGAAACTGGTATTCCCG

attggccttgttccccagggtggagctgtcacaaaatagagtgggaactgtctggcttt
cagcccaagagaatctgcatggcaagttgcattaacaaccaggcatttccggcagttcc
caacatttctgggaattttctcatccaaacgactgaaagcccactccattctcttgctt
cttactcatgctttctttgtataatggtaattatgttttaaaaaatcctgggctatgtt
gtttcatggaacaatttagaacttattggtcaaactctgaagcaaaggtatataaaagg
tagttagagatgtttagggaatattcaaagcacatttttgggtcactcataattgatct
ttatattcatatatgtatatatatatatataacataatgtacccatcttaacatatcaa
agctaaaccagtattaaaaacaactgactatggtctattgatacaatatatgatgccca
agtacactcttcattgctactgcatatctaaaatcatttatttatttatccatcCATCA
AGAGTGTATTGAGAGCCTGACAACATACC[ag]GCATCAAGCCCTGGAGGTCTTTTTAA
GGCTGAGCcaatatagctatggataacattctaaaactgatagcatattttcatgtttt
atagtctttccacagactagttcaaaatgaacactgcctgagaggggctttaagatgac
tgactagaggtactggacacctgtttccccagcaaagaagagccaaaatagcaagtaga
taatcatactttgaatagacatctaagagagaatgctggaattcagcagagaagtgaca
gaaaacacctgagatactgaaggagagggaggcaaggtagacagcctggctggaatcag
ctgggagcccagagagggtccctagtgagaggaaagggtaagtgagagattcccagtgg
tacatgttcccatgttgactgctgaaatcctagtcataagagtctctcaaaccccaagg
accctgaaactggtattcccg

FIGURE 23vvvv

54915
GTTTAAATTTGCCGATTAGTTTCGATGATTCACCAGTGCTTGATGATTAAGGGGTATTG
GTGCAGTGCCACTGAGTTGCTGTTCATAGTCTCCAGTAAGGGCAGTACAAGAGAGGAGA
AAAGTAAAGTTGCACATCAGGCCAATACATTTCCATGTCCCTACAGCCCATGGGTATTT
TTCTCTGAAGTTTAAAATTACAGCTCAAGAAGATCATATGTATTTATGTAATCTGCCTT
TAACCAGGCCACCTTGCTTCCCTAATGCTGTTGTTTTTTTCCCTTCGTTTATTTATCTT
TAATTGACACCTGTTGCTATTCTTATGCCTGCTCACCTTCACATAAATGTCAGCATCCA
TGCACCATGTATGTCACACACACACACACACACACACACACACACACACACACCCCTCT
AAAAGTTCTGATGAGTATTTGATAAATAGTAGAGTTTTGAGGAGAGATGGAGGAAAGTG
TTTACAAGTTTAACTTTTTGAATTTGCTTTTAACTCTCTGCTGTTCCCTCACCTGTAAA
ATCTGCCTCATCTCTGCCCCTCTTTCTTC[tc]TGCAAACCTCACTTCTCATAGCCTCC
TCCAGCAGCACTGACTTCTGGAGATTCCCTGTCAGTGAAATAAAACTGGAAAGCTGGTC
TCATAATAAAAGCCCAACAGTTTATGGCAAAGCCCAACCACCTGTGGTTCTTCAGGTG
TGGTTTTCTTGAGGAGTGCTTATTTACCCTGCCACATTTTCCTCTCTTTCTCTCCAAGG
AGGCTTTCTCTCCAGGGTGGATTAAGTGAAATTATGCTGTTACTTAGGGACTGATTTAC
ATATTTCTTATCCCTCACACTCTGGGTTTCTCTATGTTAGCTACATCTAGGAAAAAAAT
GGGGAAAAAAATCACCTTGATTGGAAGTGCAGTTAATTCCTGAAAATAAAGCCTGATCA
CGAGTGGTAATCACAGATCAATTAGTTACTGGATCCCTAGATAATGCATCCCTGTCATT
GTGAGACAAAAGAGGGGAAAG

gtttaaatttgccgattagtttcgatgattcaccagtgcttgatgattaaggggtattg
gtgcagtgccactgagttgctgttcatagtctccagtaagggcagtacaagagaggaga
aaagtaaagttgcacatcaggccaatacatttccatgtccctacagcccatgggtattt
ttctctgaagtttaaaattacagctcaagaagatcatatgtatttatgtaatctgcctt
taaccaggccaccttgcttccctaatgctgttgtttttttcccttcgtttatttatctt
taattgacacctgttgctattcttatgcctgctcaccttcacataaatgtcagcatcca
tgcaccatgtatgtcacacacacacacacacacacacacacacacacacacacccctct
aaaagttctgatgagtatttgataaatagtagagttttgaggagagatggaggaaagtg
tttacaagtttaactttttgaatttgcttttaactctctgctgttccctcacctgTAAA
ATCTGCCTCATCTCTGCCCCTCTTTCTTC[tc]TGCAAACCTCACTTCTCATAGCCTCC
TCCAGCAgcactgacttctggagattccctgtcagtgaaataaaactggaaagctggtc
tcataataaaagcccaacagtttatggcaaagcccaaccacctgtggttcttcaggtg
tggttttcttgaggagtgcttatttaccctgccacattttcctctctttctctccaagg
aggctttctctccagggtggattaagtgaaattatgctgttacttagggactgatttac
atatttcttatccctcacactctgggtttctctatgttagctacatctaggaaaaaaat
ggggaaaaaaatcaccttgattggaagtgcagttaattcctgaaaataaagcctgatca
cgagtggtaatcacagatcaattagttactggatccctagataatgcatccctgtcatt
gtgagacaaaagaggggaaag

FIGURE 23wwww

54918
CCCAAAATCTTAGGATGCTGCCTTAAACATCATGGTAGAATAATGTAACTAGCTACCCA
CGATTTCCTTCTTTAATTCATTTTGTGTTTTATCTCCCCAGGAAAGTATTTCAAGCCTA
AACCTTTGGGTGAAAAGAACTCTTGAAGTCATGATTGCTTCACAGTTTCTCTCAGCTCT
CACTTTGGGTAAGTCAGTGCCATTAGACCAAGATTTCTCATTCTCGCACTATAGATATT
TCAGACTGAAATATCCTTGCTTGTCTGGGGCTGTCCTGCACAGGATATCTGGCAGCATC
CTTGACCTCTACCTGCAATGTGTTCTTCCCTGGCTTGGGGTCATTTACTTTACCTCTT
GGTGTCTCCCTTTCCTTAAGTGTAAAGTGTGGATCATAATGACCTATTTCCCAGATGCA
TTGTGAGGATTCAATAGCATGGTTCATGGAAAGTACCTCATACAGTGCTTCTTGGTGCA
TACTAAGTGCTCAATAAAGCTTAGTTATTCTGATTATTATTCTACTACAAAATGGGTAT
ACTATAATGTTGTGAGTGAGTGTGGATAA [ga] GTACCTAGTGGGTGGCAGTCACAAAA
GAGATAAACAATAAGTCGCTGTTTCTTCATACGTACTTCTTACTTTTGAAAAGATGAGA
AAAGTCTGGGCCATGTCACAAACATTGCCAAAAATAAGACAATAAAAAGCACAGTTGTC
AGAGTTAAACCACAACAGTACCAAACTCTACCATTTCTTTTCTTTTTCTCCCACTAGTG
CTTCTCATTAAAGAGAGTGGAGCCTGGTCTTACAACACCTCCACGGAAGCTATGACTTA
TGATGAGGCCAGTGCTTATTGTCAGCAAAGGTACACACACCTGGTTGCAATTCAAAACA
AAGAAGAGATTGAGTACCTAAACTCCATATTGAGCTATTCACCAAGTTATTACTGGATT
GGAATCAGAAAAGTCAACAATGTGTGGGTCTGGGTAGGAACCCAGAAACCTCTGACAGA
AGAAGCCAAGAACTGGGCTCC

cccaaaatcttaggatgctgccttaaacatcatggtagaataatgtaactagctaccca
cgatttccttctttaattcattttgtgttttatctccccaggaaagtatttcaagccta
aacctttgggtgaaaagaactcttgaagtcatgattgcttcacagtttctctcagctct
cactttgggtaagtcagtgccattagaccaagatttctcattctcgcactatagatatt
tcagactgaaatatccttgcttgtctggggctgtcctgcacaggatatctggcagcatc
cttgacctctacctgcaatgtgttcttccctggcttggggtcatttactttacctctt
ggtgtctccctttccttaagtgtaaagtgtggatcataatgacctatttcccagatgca
ttgtgaggattcaatagcatggttcatggaaagtacctcatacagtgcttcttggtgca
tactaagtgctcaataaagcttagttattctgattattattctACTACAAAATGGGTAT
ACTATAATGTTGTGAGTGAGTGTGGATAA [ga] GTACCTAGTGGGTGGCAGTCACAAAA
GAGATAAACAATAAGTCGCtgtttcttcatacgtacttcttacttttgaaaagatgaga
aaagtctgggccatgtcacaaacattgccaaaaataagacaataaaaagcacagttgtc
agagttaaaccacaacagtaccaaactctaccatttcttttctttttctcccactagtg
cttctcattaaagagagtggagcctggtcttacaacacctccacggaagctatgactta
tgatgaggccagtgcttattgtcagcaaaggtacacacacctggttgcaattcaaaaca
aagaagagattgagtacctaaactccatattgagctattcaccaagttattactggatt
ggaatcagaaaagtcaacaatgtgtgggtctgggtaggaacccagaaacctctgacaga
agaagccaagaactgggctcc

FIGURE 23xxxx

54966
GCAGGACTGCAGACATGACTCATGGCAGGGTAGCTGCTGAGGCACGTCCCATCTCCTTT
CAGTTCAGGAGAGGCTGTGGGAAGAGGGAAGAACTGAGCACACATGAAGATTTGGCAGA
GGGAGGAGGCCAAGTAGGGAGGAAGTGGAATAATTGATATTGGAGCCAGACATATAATC
AGATGAAACCTGGGCAAAACCAAACGAGGTCCAGACATAAGGAGAAGGAGAGCAGGCGA
AAAGGCAATAGAGATCTGTGGCATGAGATAATCCTATGTCCGTGGGATTTTCCCATGGA
TGGTACAACTGGCACAGGACGATGTTATTCCTCCCCTCTGGTGAAACCAATATGGCAGC
AGAAGGCAGGGAGGGTGGGGAGGAGGGTGTAGTTTGTCTGCACAAGCATCATCAGCATA
TTTTCAGGAGCTTCTGAGAGCTGATGAAGGATCATTTGCTGCAGATACTTTATATTCAC
TCGGTCAGCCAACTTGTATTGAGCAATTGCTGGGGCACAGCAGTGAGTGAGGTGCGCTA
CAGAAACACAGTTGAAAAGAATCTGACTT[tc]GCCCTCAATGAACCTGCAGTCAAGTT
AGAAGCACAGAGGTCAACAGACAAATAAGATAAAGGCATTAGTTTCTGTACTGGAGCAT
AACACCAATACTGCCATTGCTCAGAATGTTTCTAGAACCCCTAAAAGTTCAGAACTGTC
TTCAGCATCATTTCAGGAGCCAGACAAGAAAACCAGTCTCATTTCTTTATTGTCATGAC
CTGGGTTTGACCAGAAACAATATTACTCACTTGGAGCACCTCACTCCTCAGATCTGGCT
CTAGTTCTAAATATCAAACCATTCTCAAATAGCAAAGCTTTGTCACCTCCCTATACATA
TCTCATTTAAATATGTAAAGGATCTGTAGGCAATTCCAAAAAGAAGGCTCTAAAAATAT
TTAAAAAGCAATGGTCGTACCTTATAGTTTTACCTTATAGTGTATATCAATAATAGCCT
TGTAATTAAAAAACAATCATC

gcaggactgcagacatgactcatggcagggtagctgctgaggcacgtcccatctccttt
cagttcaggagaggctgtgggaagagggaagaactgagcacacatgaagatttggcaga
gggaggaggccaagtagggaggaagtggaataattgatattggagccagacatataatc
agatgaaacctgggcaaaaccaaacgaggtccagacataaggagaaggagagcaggcga
aaaggcaatagagatctgtggcatgagataatcctatgtccgtgggattttcccatgga
tggtacaactggcacaggacgatgttattcctcccctctggtgaaaccaatatggcagc
agaaggcagggagggtggggaggagggtgtagtttgtctgcacaagcatcatcagcata
ttttcaggagcttctgagagctgatgaaggatcatttgctgcagatactttatattcac
tcggtcagccaacttgtattgagcaattgctggggcacagcagtgagtgaggTGCGCTA
CAGAAACACAGTTGAAAAGAATCTGACTT[tc]GCCCTCAATGAACCTGCAGTCAAGTT
AGAAGCACAGaggtcaacagacaaataagataaaggcattagtttctgtactggagcat
aacaccaatactgccattgctcagaatgtttctagaacccctaaaagttcagaactgtc
ttcagcatcatttcaggagccagacaagaaaaccagtctcatttctttattgtcatgac
ctgggtttgaccagaaacaatattactcacttggagcacctcactcctcagatctggct
ctagttctaaatatcaaaccattctcaaatagcaaagctttgtcacctccctatacata
tctcatttaaatatgtaaaggatctgtaggcaattccaaaaagaaggctctaaaaatat
ttaaaaagcaatggtcgtaccttatagttttaccttatagtgtatatcaataatagcct
tgtaattaaaaaacaatcatc

FIGURE 23yyyy

55560
CAGGAAGTTGTTGGTGTTTGGATGGATGAATGGACTAATGGATGGATGAATAATAGATA
GATGGATTGTTGAGAGAGACAGAGAAGAGAAAAGCCTTGCCCCCAAAAGCTCACAGACT
ACTTGGAGAGAGAAGAAAGCTACCTGGAGGGAGAACCAGATGCATGAAGCAGTGCAGAT
GTGGTGCCTAATGAGTGTGTAGTCTGGAAGGGCAGCAAAAGTCGAGTGGAGTGAGAGGT
TCCTGTGTCCTGGAGCACTGAGTAGAGACTCCCTCATGGGGGTGAATCTTAAAGGATAA
AGGGGCCTCTATAATGAAAAGGAGGAGGATGGGATTTCTGGTAGAGGAAATTGCTTGAG
CAAAACCTCCAAGGTTGGAATGACTATGGTGTGTTCAGGGATGTTAGCAGACCCAGATG
GGTGGAGCGTTGAGTGTGTGTGTGTAGGAAGGAAGAGGGGAGGTGGCTGGATGAGCACA
GTGAGACCTGATTTGATTGAGAGCCTTGAACGCCACGCTGAATAATGGAGGCAATGGGA
CGCCATAGAGGGCTTTTGAGTAGACATAT[ag]TCAGTGTAGAAGGGTGAATTTCAGAT
TTTTAGACAGAATAGAGTAAGGAGAGGAGCTCTTAGAAATCATCTAGTCCAGGGCTTGT
GGCAGAGCCCTGAGGTTTTAAGAAGGCATGTCAGGGGCTACCATGACAGGCACGGAGAG
GCTGAGTGAATTGGGGTTCTTGCCACAATTCCCTTGCCTGAGATTCAACAAGAGCAGCT
GTATTACAATCTGTGCAAAATGTCATTAGGAGAAACTAGTTAGTAGCTGGGCGTGGTGG
CATGCAACTGTTGTCCCAGCTACTCGGGAGGCTGAGGCCGGAGAATCGCTTGAAGCTGG
GAGGCGGAGGTTGCAGTGAGCAGAGACTGTGCCACTGCACTCCAGCCTGGATGACAGAG
CAAGACTCTGTTTCAAAAAAAAAAAAAAAAAAAAACTAGTCAGGACTCTTTCAGATACAA
GTAATAGAAACCAACTCAAAC

caggaagttgttggtgtttggatggatgaatggactaatggatggatgaataatagata
gatggattgttgagagagacagagaagagaaaagccttgcccccaaaagctcacagact
acttggagagagaagaaagctacctggagggagaaccagatgcatgaagcagtgcagat
gtggtgcctaatgagtgtgtagtctggaagggcagcaaaagtcgagtggagtgagaggt
tcctgtgtcctggagcactgagtagagactccctcatgggggtgaatcttaaaggataa
aggggcctctataatgaaaaggaggaggatgggatttctggtagaggaaattgcttgag
caaaacctccaaggttggaatgactatggtgtgttcagggatgttagcagacccagatg
ggtggagcgttgagtgtgtgtgtgtaggaaggaagaggggaggtggctggatgagcaca
gtgagacctgatttgattgagagccttgaacgcCACGCTGAATAATGGAGGCAATGGGA
CGCCATAGAGGGCTTTTGAGTAGACATAT[ag]TCAGTGTAGAAGGGTGAATTTCAGAT
TTTTAGACAGAATAGAGTAAGGAGAGGAGctcttagaaatcatctagtccagggcttgt
ggcagagccctgaggttttaagaaggcatgtcaggggctaccatgacaggcacggagag
gctgagtgaattggggttcttgccacaattcccttgcctgagattcaacaagagcagct
gtattacaatctgtgcaaaatgtcattaggagaaactagttagtagctgggcgtggtgg
catgcaactgttgtcccagctactcgggaggctgaggccggagaatcgcttgaagctgg
gaggcggaggttgcagtgagcagagactgtgccactgcactccagcctggatgacagag
caagactctgtttcaaaaaaaaaaaaaaaaaaaaactagtcaggactctttcagatacaa
gtaatagaaaccaactcaaac

FIGURE 23zzzz

56733
TACCAAAGGGCAAGTAGGGAAACAGACCAACAGAGATGTTACCTTCTGAATAATTGGAC
CCAGGAAGAGGAGTGTAACCTAAGAGAGGAAGATACTTGATTATACCAGTCTTTGTGGA
TGAAAATATCTAGCAGTATTCATAGCAAATGCAGTAGGAAGGAGAGAGTTAATCACAAA
CAGAAAGTAAGCAGAGAGTGGGACCAAGAGTGGGGATGGGAGTTCAGCGAGTCACTCAC
TAGAGTGGCCAGCTCTCCGCCAGCTGATCACACCAAGAGAGAAGATGATGAGGCCCAGG
CCCAGAGTCACTGCAGACACAGAAACCTTCAGGGTCTGCATGGGGGACAGCCCAGGTGC
TGCAAAAAATAGAAACTTACTTGACCCAGTTTCTGTTGCTCACCCCCAGGGCAATTCCA
TTTATTGCAGCCACCTCTCAGTGGGTTAAAAGGTCCTTTATCCCAGCTCCAAGGGTCTA
GCTCACACCACCCACTCCCAAGAAAATGATCTTTCTCAAATCAAACCCTCGTCCCATGG
ACCTCTACTCCTAGAGTAAGCCTGGGGAA [ag] CCATCTCCCCAGAATTAGCATCCTGG
CTTCCAGGTCCTCTCTAATACAGTGGGGCCTCTCAAGGCATCCTCTTTCCTTCCTTTAC
CTCAAAGCCACCCTTATCAGGATAAAGGGCTCCTCACTGTCCTCTCCATTGCCCCCACG
GTAACAATGTTTGCTTCCTTACTTTCTCCAACTGAGCAGCTTCCTATTACACTGTCTTA
CCACATGTCTTAACCTCCAGTGGATCCATCCTGTGAGTTATCCTACTACTTGTGTACCT
TCTACATCTAGATCTCCCATGTGTCCTTTCAGAGCTTGTCTCCATCCCACTCCACAGCC
CCTGCACTTCCTTGGGCCGGTCCTGTTCTGAATCATGTCCCACTCAGATTCTTTTCCCA
TGATAAAATGAACACTCCATTTCTAAAGGGAGGCTCTTGTGCACGCTGTGAGGAGACGT
TCCCCAGGAAAGTTCAAGTGA

taccaaagggcaagtagggaaacagaccaacagagatgttaccttctgaataattggac
ccaggaagaggagtgtaacctaagagaggaagatacttgattataccagtctttgtgga
tgaaaatatctagcagtattcatagcaaatgcagtaggaaggagagagttaatcacaaa
cagaaagtaagcagagagtgggaccaagagtggggatgggagttcagcgagtcactcac
tagagtggccagctctccgccagctgatcacaccaagagagaagatgatgaggcccagg
cccagagtcactgcagacacagaaaccttcagggtctgcatgggggacagcccaggtgc
tgcaaaaaatagaaacttacttgacccagtttctgttgctcacccccagggcaattcca
tttattgcagccacctctcagtgggttaaaaggtcctttatcccagctccaagggtcta
gctcacaccacccactcccaagaaaatgatctttctcaaatcaaaccctcgtcccATGG
ACCTCTACTCCTAGAGTAAGCCTGGGGAA [ag] CCATCTCCCCAGAATTAGCATCCTGG
CTTCCAGgtcctctctaatacagtggggcctctcaaggcatcctctttccttcctttac
ctcaaagccacccttatcaggataaagggctcctcactgtcctctccattgcccccacg
gtaacaatgtttgcttccttactttctccaactgagcagcttcctattacactgtctta
ccacatgtcttaacctccagtggatccatcctgtgagttatcctactacttgtgtacct
tctacatctagatctcccatgtgtcctttcagagcttgtctccatcccactccacagcc
cctgcacttccttgggccggtcctgttctgaatcatgtcccactcagattcttttccca
tgataaaatgaacactccatttctaaagggaggctcttgtgcacgctgtgaggagacgt
tccccaggaaagttcaagtga

FIGURE 23aaaaa

66854
CAAAAGGTCACCCCACAGTCCCCACTCCAAGGCAGGTTGATAGCAGGGATCTCAGGGTG
CCCATGGATCAAGGACTAAGTCAGAGTCGGGGTCCCTCAGGCCGAGGGTAACGTAGGTG
GTGCCTGCCAGGCTCTCCTCGCCCAGGGGGGCTGAGAATGTCTAAACCCGGGTGGCTGT
GACCCCTAGGCAGAGCCAGCCCAGCCCTTGCCAGGGATGGAGACCGGCCTCGAGGAGGC
CAAGCCCTGGGGGTCCACAGGCCTGTGGGCTTCGGGGAGGCTCTGCTCCCTGTGGCCCT
GTGTGGCCCAGGCTGCTGAGTCATCAGAACCTCGGGGGCGCCGCGGGCCCCACATTCCG
CCCAGGCCTCTCTCTGACCCCCTTCCCAGCCCATCTGTGTTTTTGGAAAACAGAGCCAG
AGCCCCCCGCGGCCCTGCCAGCTTGCGGCTGCTCACGCTGGGACTCAAATCGCACCCTT
CTGTCTTCAAAGTCCACCTTCACTTCAAAGCTCGGTCCCACCCCAGCCCGGCCTCCACA
GGGCCACCACCTGCCCACACCCAGGCCCGCTGCTGCCCAGTTTCGGAGGGACCTTGGGC
ATCCCCTGATCCTCTCTAGAGCG[ct]GGGGTTCCTGGCATGGGCCCGTTACACATGGG
TGGCTCGGTGGGTGGTGAGGACGGGGCTGGGAGAAGATCCTGGGGACCCCATGGTGGAG
GCAATGAGGCACCCAAACCCCAACTCCAGCGATGGCTGCTTCCACGGGGCCCTCCGAGC
CCTGACCTTCAAGGTGCAAGAAAAGCTTTCAGGGGCAGGGGTGAGTGGAAGGTGGGCTT
CCTCCCTTGCCACCTGGGGGGCGGGCCCAGGACAGATGCTCCGTGAGAGCACTTCCCAA
CCTAGGCCCAGCTGTGGGGAAGGAGGGAGCAGGCGGCTGGGCTCCAGGCAGGGGGAAGA
GTTGCCTGAGAACTCAGGGAGAGAGGGAGGGCTGGGGCACCCCATGCCAGCTCCAGCTG
CAGCACCAGAGCTCAGAGCAG

caaaaggtcaccccacagtccccactccaaggcaggttgatagcagggatctcagggtg
cccatggatcaaggactaagtcagagtcggggtccctcaggccgagggtaacgtaggtg
gtgcctgccaggctctcctcgcccaggggggctgagaatgtctaaacccgggtggctgt
gacccctaggcagagccagcccagcccttgccagggatggagaccggcctcgaggaggc
caagccctgggggtccacaggcctgtgggcttcggggaggctctgctccctgtggccct
gtgtggcccaggctgctgagtcatcagaacctcgggggcgccgcgggccccacattccg
cccaggcctctctctgacccccttcccagcccatctgtgtttttggaaaacagagccag
agccccccgcggccctgccagcttgcggctgctcacgctgggactcaaatcgcaccctt
ctgtcttcaaagtccaccttcacttcaaagctcggtcccaccccagcccggcctccaca
gggccaccacctgcccacacccaggcccgctgctgcccagtttcggagggaccttgggc
atCCCCTGATCCTCTCTAGAGCG[ct]GGGGTTCCTGGCATGGGCCCGttacacatggg
tggctcggtgggtggtgaggacggggctgggagaagatcctggggaccccatggtggag
gcaatgaggcacccaaaccccaactccagcgatggctgcttccacggggccctccgagc
cctgaccttcaaggtgcaagaaaagctttcaggggcaggggtgagtggaaggtgggctt
cctcccttgccacctggggggcgggcccaggacagatgctccgtgagagcacttcccaa
cctaggcccagctgtggggaaggagggagcaggcggctgggctccaggcagggggaaga
gttgcctgagaactcagggagagagggagggctggggcaccccatgccagctccagctg
cagcaccagagctcagagcag

FIGURE 23bbbbb

66865
ATTCCCTGACCAGGGCCCTGGGACCCACCGCACAGCTGAGCTGGCCCGAGCTGAAGAGT
TGTTGGAGCAGCAGCTGGAGCTGTACCAGGCCCTCCTTGAAGGGCAGGAGGGAGCCTGG
GAGGCCCAAGCCCTGGTGCTCAAGATCCAGAAGCTGAAGGAACAGATGAGGAGGCACCA
AGAGAGCCTTGGAGGAGGTGCCTAAGTTTCCCCCAGTGCCCACAGCACCCTCCGGCACT
GAAAATACACGCACCACCCACCAGGAGCCTTGGGATCATAAACACCCCAGCGTCTTCCC
AGGCCAGAGAAAGTGGAAGAGACCACAAACCGCAGGCAATTGGCAGGCAGTGGGGGAGC
CAGGGCTCTGCAGTCTTAGTCCCATTCCCCTTTGATCTCACAGCAGGCAGGGCACCCAG
GCCTTATAGGAATTCACCCTGGACCATGCCCTAAAATAACCTCACCCCAAATACAATAA
AGGGACGAAGCACTTATAGATACCACAGACACATGTGTTTCATTTTTAGTTTTGTTAAA
AAAAAATTCTGACAAATCAGAAATGGGGGTTCAGGAGTGGTGGTGATGCAAAAGATGGA
AGCCATGGGGTGGGGGCTGTCAGGGGTGGGGGCAGTAGTGTCTCCTT[ct]ACCCCCAC
CCTGGTGTCCTCTCCTGAAGGACAGACGGTCACATTCCAAAATGGGCGAGTCTTCTACC
GTGTCTGTTCAACTGAGAAGAAAACGTAGCATGGTCAGAATAAGGCATGAAAAGGGGAA
AGTGAGGCAGGAACACACGGCACACATGCAGACACTGGTGTACTGCCTGGGTTCAGAGG
ACGGACGTGGGGGTGAGGGAAGGGATGTAATATGATGAGAGAAGACAGAAACCCCACAT
AAAGGTCAGAAAAACATCCCAACACAGCATCAAAGACCAGGGGGCATGAACCAGTCAAG
TGTCCATTATGCATCAGATGCCCATGACCTATGTGATGGGATTTAGGACAAACACACTA
AGGAACAGGGAGGACCTAAAG

attccctgaccagggccctgggacccaccgcacagctgagctggcccgagctgaagagt
tgttggagcagcagctggagctgtaccaggccctccttgaagggcaggagggagcctgg
gaggcccaagccctggtgctcaagatccagaagctgaaggaacagatgaggaggcacca
agagagccttggaggaggtgcctaagtttcccccagtgcccacagcaccctccggcact
gaaaatacacgcaccacccaccaggagccttgggatcataaacaccccagcgtcttccc
aggccagagaaagtggaagagaccacaaaccgcaggcaattggcaggcagtgggggagc
cagggctctgcagtcttagtcccattcccctttgatctcacagcaggcagggcacccag
gccttataggaattcaccctggaccatgccctaaaataacctcaccccaaatacaataa
agggacgaagcacttatagataccacagacacatgtgtttcatttttagttttgttaaa
aaaaaattctgacaaatcagaaatgggggttcaggagtggtggtgatgcaaaagatgga
agccatggggtgggggctgtcagGGGTGGGGGCAGTAGTGTCTCCTT[ct]ACCCCCAC
CCTGGTGTCCTCTCCTgaaggacagacggtcacattccaaaatgggcgagtcttctacc
gtgtctgttcaactgagaagaaaacgtagcatggtcagaataaggcatgaaaaggggaa
agtgaggcaggaacacacggcacacatgcagacactggtgtactgcctgggttcagagg
acggacgtgggggtgagggaagggatgtaatatgatgagagaagacagaaaccccacat
aaaggtcagaaaaacatcccaacacagcatcaaagaccagggggcatgaaccagtcaag
tgtccattatgcatcagatgcccatgacctatgtgatgggatttaggacaaacacacta
aggaacagggaggacctaaag

FIGURE 23ccccc

66902
CTCTGAAGGCTTGCCTGGTGCTCACTCAGCCCGTGAAGAGGGCCTGCTGGTCCTCTGGA
GCCCACAGCCCTTTGTCCAGAGGCGACTCCTAACCTTTAGCAGGCTCTGCCCTAACTTA
CAGTCCCACCATTGTCTGCCCCACATCCTGTCTGCCTGTCTGTGCTCCATTCTGGCCCA
TCCTAGGTGTCTCTGGCTGCAAAGCCTTTCCTGGGCTCAGCCTTCTGCCTTGAACGGGC
CCTGACCATGAGTCCCCATGTGCCCAGCCCATACCTTTTCCCTGTCCAGCCAGGAGCCA
ACACAGGCCTGGAGCATTGCCTGTGGTATGGCCTGCTCGCTGCTGTTCCCGGCCTGGGT
GGTCACGGACATGCAGAGGTGGCACTCAGAGTCTCGCGGCAGCCATTCTCCTGTCGGCG
ACCCTGGAGATGTGAGCATTAGGGGGAAAGCAGGCAAGGCCACCCTACAGAGGTGTTTG
GTTTCTGTCCTCCTTGGTGCATTGCAGTGGGACCACAGAGGGAGAGGGTCATGCAGTGG
CAGGGTAGGGGGAGGAGGAGAGCAGGCATTGGGCTAAGGAG[gt]GGGCAGTGGGCTCA
CTTGGGCCAGCGCTGTCATCCATGGAGCACCGGAGGACGAGGCGGCAGACCAGCTGGGG
CAGCATGCGGCCCAGCAGCGTGTCGAGCAGGATGACGGAGTAGCGCTCAGCCAGGCACT
GGCAGATGCCGCCCGCCACCAGAGGTACCACGCGGCACACCTGGGCCACTGCCACAGCT
AGCGCACCCTGGGGCGGGGGCGGAGAGAGGCCAGCATGGGACCTTCACTTGGCAAGCCT
CCACTCTCTGCCCAGCACCCAGCTGGGCACTTCCTACGCATTCCCTCATTCTCTTCTAG
AAGGGAGGGCAAGGCTATTCACAAATAAGGACACTGGGGATCAGAGAGTCCAGGGGATG
CAGGGGACTCACACAGGGTCACTGAGTGTAGGAGCCAGCTTCAGACCTACGTCTGGCCC
CAAAGGCTCTGGCCCACAGCT

ctctgaaggcttgcctggtgctcactcagcccgtgaagagggcctgctggtcctctgga
gcccacagccctttgtccagaggcgactcctaacctttagcaggctctgccctaactta
cagtcccaccattgtctgccccacatcctgtctgcctgtctgtgctccattctggccca
tcctaggtgtctctggctgcaaagcctttcctgggctcagccttctgccttgaacgggc
cctgaccatgagtccccatgtgcccagcccataccttttccctgtccagccaggagcca
acacaggcctggagcattgcctgtggtatggcctgctcgctgctgttcccggcctgggt
ggtcacggacatgcagaggtggcactcagagtctcgcggcagccattctcctgtcggcg
accctggagatgtgagcattagggggaaagcaggcaaggccaccctacagaggtgtttg
gtttctgtcctccttggtgcattgcagtgggaccacagagggagagggtcatgcagtgg
cagggtagggggaggaggAGAGCAGGCATTGGGCTAAGGAG[gt]GGGCAGTGGGCTCA
CTTGGGCCAgcgctgtcatccatggagcaccggaggacgaggcggcagaccagctgggg
cagcatgcggcccagcagcgtgtcgagcaggatgacggagtagcgctcagccaggcact
ggcagatgccgcccgccaccagaggtaccacgcggcacacctgggccactgccacagct
agcgcaccctggggcgggggcggagagaggccagcatgggaccttcacttggcaagcct
ccactctctgcccagcacccagctgggcacttcctacgcattccctcattctcttctag
aagggagggcaaggctattcacaaataaggacactggggatcagagagtccaggggatg
caggggactcacacagggtcactgagtgtaggagccagcttcagacctacgtctggccc
caaaggctctggcccacagct

FIGURE 23ddddd

66934
GGAGAGCAGCAGCTGGAGGGCAGGCTGGGAGCGCTTGTGAGGGAGAGGAGCTATGGACG
TCTGCTTCTCTGCCAAGGGAGAGAGTGAGGTAGGCCTGGGCCCGCTGACTTCAGGGTGA
GGCCACAGCTACTGCAGCGCTTTTTATTTATTTATTTATTTACTGAGATGGAGTCTTGC
TCTGTCACCCAGGCTGGAGTGCAGTGGTGCAATCTCGGCTCACTGCAACCTCTGCCTCC
TGGGCTGCAGTGATTCTCCTGCGTTCAAGTAATTCTCCTGCCTCGGCCTTCTGAGTAGT
TGGGATTACAGGCATATGCCACCACACTTGGCTAATTTTTTGTATTTTTAGTAGAAATG
GGGTTTCACCATGTTGGCGAGGCTGGTCTCGAACTCCTGACCTCAAGGATCCTCCTGCC
TCGGCCTCCTAAGGTGCTGGGATTGCAGGTGTGAGCCACCACGTCTGGCCATACTGCAG
CACTTTAAAGGACGGTGTCTTTTTCTTTCTCATAAAAGAGAATAGGACTTTATTAGCA[
tc]TGGTGCAGACATTGTATTACACAGGAATGGGTCCCTAGCTTGCACAACCCCAGCTG
AGCTTTCAGCAGATAAATCACAGCAGAAATAGAATCACCCTAGGACTTTCAATCAAAAG
CTGGAAGTCCACCTTACAGAAAGACAAAAAGAAACCCCTTTTTATATCTTAACAAAGCA
ATAGCTCTCAAGCAGCAGAGCATCTCGAGGAAGAAAGCTTGCCCGGTCGCCATCCCATC
ATGCCAGAGCGTGCAGTGTCCACCCTTGACTACGCTGGGGAATTGCTGATTTTTTGAAA
AAGCTTAACTTAACAATTTCTGATGTCTATCTTTTAGAGTTCTGTATGTTCCCATTTTT
TATTCTTCTGAATTTTGAATTGCAAGTAGCTGTAAAATCCAATCTTTGAGTGCATGGGG
GTGGGTGTGAGGCGGGGCTCAGCTTCAACCCCCTGTCCTGTAAAGCAGTGGCTGGTTTT
TCCTGAGCCCAGCCCTGGGAG

ggagagcagcagctggagggcaggctgggagcgcttgtgagggagaggagctatggacg
tctgcttctctgccaagggagagagtgaggtaggcctgggcccgctgacttcagggtga
ggccacagctactgcagcgctttttatttatttatttatttactgagatggagtcttgc
tctgtcacccaggctggagtgcagtggtgcaatctcggctcactgcaacctctgcctcc
tgggctgcagtgattctcctgcgttcaagtaattctcctgcctcggccttctgagtagt
tgggattacaggcatatgccaccacacttggctaattttttgtatttttagtagaaatg
gggtttcaccatgttggcgaggctggtctcgaactcctgacctcaaggatcctcctgcc
tcggcctcctaaggtgctgggattgcaggtgtgagccaccacgtctggccatactgcag
cactttaaaggacggtgtctttTTCTTTCTCATAAAAGAGAATAGGACTTTATTAGCA[
tc]TGGTGCAGACATTGTATTACACAGGAATGGGTCCCTagcttgcacaaccccagctg
agctttcagcagataaatcacagcagaaatagaatcaccctaggactttcaatcaaaag
ctggaagtccaccttacagaaagacaaaaagaaaccccttttatatcttaacaaagca
atagctctcaagcagcagagcatctcgaggaagaaagcttgcccggtcgccatcccatc
atgccagagcgtgcagtgtccacccttgactacgctggggaattgctgattttttgaaa
aagcttaacttaacaatttctgatgtctatcttttagagttctgtatgttcccattttt
tattcttctgaattttgaattgcaagtagctgtaaaatccaatctttgagtgcatgggg
gtgggtgtgaggcggggctcagcttcaacccccctgtcctgtaaagcagtggctggtttt
tcctgagcccagccctgggag

FIGURE 23eeeee

66938
CAGCCATGGTTCGCGGTGCCCTCGGCTGCCCTGGGCCAGAGCTGGGGCTAGCTTTCACC
TTGTTGAGACCCAGGACTCTGTCCCCCAAGCCTGTCTTCGCCAGCGCCTTGACCCCACC
CCTCATATACTGTGTCCTGGAAAACGTGGACACGGGAGACCACAGCCAGGGCGAGGTAT
CGCCCCTCCATCCCCCCAGGCCCAATGAGAAGCAGTTGGCCAAGGTGATCCAGGTGGCA
GAGGCAGCATCAGACCCAGTCTCCTGTCAGGCACCACCTTGGGTGCCGGTCCCCAGATG
CCCTGGCGGGGAGTGTGCATGCTCCCGGAGCCCCCAGGTCACCCCATGTGAGCCAGGCC
CACAGAGCTTGGCTCTGCAATGCCTGCTGGGCTGCTGCCCATGCTCCACCCCTTCTGGG
AAGCTAAAAGACAGCCCTTCAGTGTCCAGAGACCTGCCTGGCCTTGGAGCCTGGGTTTC
ACATGCCCACCGGGCTGGCAGGGGCACTCAGCTGCCTCCAGCCCCGGCGGTCACCCTGG
CATTGGGTCCATCTAACTGCTCCCCAGTCACAAGGCAGCTGCTCCCCAAGTCTCCCCAA
A[ct]CTGCTGGCCCCTCTAGAAGCCTCTGTCCATTCCTGGAGGACCGAGGGCAGCCTG
CATGCCATCCCGCACACAGCCTTCTGTCTGGGCATCCTGCCTTCACACATGCTGCACAG
GGAGGAAACTCTTATACCACATTCCTTAAGCAGAGACTGAAGCCTGGAGCCAGGCACAT
GGCACATGCTCCCACCCACCCAGGACACACTGCGGTGTGGCTGCCTCCAGGCTGGCCCC
CTAGATTGCgTCTGCTCCTGGCATGGATAACTGGCGCCTTTGCCTGGCCGTTGGGGCAG
TGTTTGCCTTCCCCTGTCGGCAGCAAATATTTACTGTCCTCCGTCTCCAGGACTCTCCA
GGCCTGAGCAGACCCCGGGGGGATGAGTGTGGACTCAGCGGTGCTGAGGGTAGCCCCCT
GCCCTTCGGGTCCTGGTGCCC

cagccatggttcgcggtgccctcggctgccctgggccagagctggggctagctttcacc
ttgttgagacccaggactctgtcccccaagcctgtcttcgccagcgccttgaccccacc
cctcatatactgtgtcctggaaaacgtggacacgggagaccacagccagggcgaggtat
cgcccctccatccccccaggcccaatgagaagcagttggccaaggtgatccaggtggca
gaggcagcatcagacccagtctcctgtcaggcaccaccttgggtgccggtccccagatg
ccctggcggggagtgtgcatgctcccggagcccccaggtcaccccatgtgagccaggcc
cacagagcttggctctgcaatgcctgctgggctgctgcccatgctccaccccttctggg
aagctaaaagacagcccttcagtgtccagagacctgcctggccttggagcctgggtttc
acatgcccaccgggctggcaggggcactcagctgcctccagccccggcggtcaccctgg
cattgggtccatctaactgctccccagtcacaAGGCAGCTGCTCCCCAAGTCTCCCCAA
A[ct]CTGCTGGCCCCTCTAGAAGCCTCTGTCCattcctggaggaccgagggcagcctg
catgccatcccgcacacagccttctgtctgggcatcctgccttcacacatgctgcacag
ggaggaaactcttataccacattccttaagcagagactgaagcctggagccaggcacat
ggcacatgctcccacccacccaggacacactgcggtgtggctgcctccaggctggcccc
ctagattgcgtctgctcctggcatggataactggcgcctttgcctggccgttggggcag
tgtttgccttcccctgtcggcagcaaatatttactgtcctccgtctccaggactctcca
ggcctgagcagaccccggggggatgagtgtggactcagcggtgctgagggtagccccct
gcccttcgggtcctggtgccc

FIGURE 23fffff

66984
GGCAGAGGCAGCATCAGACCCAGTCTCCTGTCAGGCACCACCTTGGGTGCCGGTCCCCA
GATGCCCTGGCGGGGAGTGTGCATGCTCCCGGAGCCCCCAGGTCACCCCATGTGAGCCA
GGCCCACAGAGCTTGGCTCTGCAATGCCTGCTGGGCTGCTGCCCATGCTCCACCCCTTC
TGGGAAGCTAAAAGACAGCCCTTCAGTGTCCAGAGACCTGCCTGGCCTTGGAGCCTGGG
TTTCACATGCCCACCGGGCTGGCAGGGGCACTCAGCTGCCTCCAGCCCCGGCGGTCACC
CTGGCATTGGGTCCATCTAACTGCTCCCCAGTCACAAGGCAGCTGCTCCCCAAGTCTCC
CCAAACCTGCTGGCCCCTCTAGAAGCCTCTGTCCATTCCTGGAGGACCGAGGGCAGCCT
GCATGCCATCCCGCACACAGCCTTCTGTCTGGGCATCCTGCCTTCACACATGCTGCACA
GGGAGGAAACTCTTATACCACATTCCTTAAGCAGAGACTGAAGCCTGGAGCCAGGCACA
TGGCACATGCTCCCACCCACCCAGGACACACTGCGGTGTGGCTGCCTCCAGGCTGGCCC
CCTAGATTGC [ga] TCTGCTCCTGGCATGGATAACTGGCGCCTTTGCCTGGCCGTTGGG
GCAGTGTTTGCCTTCCCCTGTCGGCAGCAAATATTTACTGTCCTCCGTCTCCAGGACTC
TCCAGGCCTGAGCAGACCCCGGGGGGATGAGTGTGGACTCAGCGGTGCTGAGGGTAGCC
CCCTGCCCTTCGGGTCCTGGTGCCCAGCAGGGGTCCAGCCCAGGGAAGAGACTGAGGCC
AGGACAGGCAGTGTTTAAGCCTGAGTTTCTGGGAAAGGTAGCCCTGGGCAGAACTTGGG
CCGAACGTTGGCCAGTGTCTCTCTCCAGCCAGGCTGTGAGGTAGCTGTTTCCAGGATGG
GCACCTTTCCACACCCAGCAATGTGGCCAGGAGCCGCCATTCACGGGTGCGACCAGCAG
ATGGCATCAGAGCCTCACTTT

ggcagaggcagcatcagacccagtctcctgtcaggcaccaccttgggtgccggtcccca
gatgccctggcggggagtgtgcatgctcccggagcccccaggtcaccccatgtgagcca
ggcccacagagcttggctctgcaatgcctgctgggctgctgcccatgctccaccccttc
tgggaagctaaaagacagcccttcagtgtccagagacctgcctggccttggagcctggg
tttcacatgcccaccgggctggcaggggcactcagctgcctccagccccggcggtcacc
ctggcattgggtccatctaactgctccccagtcacaaggcagctgctccccaagtctcc
ccaaacctgctggcccctctagaagcctctgtccattcctggaggaccgagggcagcct
gcatgccatcccgcacacagccttctgtctgggcatcctgccttcacacatgctgcaca
gggaggaaactcttataccacattccttaagcagagactgaagcctggagccaggcaca
tggcacatgctcccacccacccaggacacactgcggtgtggcTGCCTCCAGGCTGGCCC
CCTAGATTGC [ga] TCTGCTCCTGGCATGGATAACTGGCGCctttgcctggccgttggg
gcagtgtttgccttcccctgtcggcagcaaatatttactgtcctccgtctccaggactc
tccaggcctgagcagaccccggggggatgagtgtggactcagcggtgctgagggtagcc
ccctgcccttcgggtcctggtgcccagcaggggtccagcccagggaagagactgaggcc
aggacaggcagtgtttaagcctgagtttctgggaaaggtagccctgggcagaacttggg
ccgaacgttggccagtgtctctctccagccaggctgtgaggtagctgtttccaggatgg
gcacctttccacacccagcaatgtggccaggagccgccattcacgggtgcgaccagcag
atggcatcagagcctcacttt

FIGURE 23ggggg

66985
AGAGACTGAGGCCAGGACAGGCAGTGTTTAAGCCTGAGTTTCTGGGAAAGGTAGCCCTG
GGCAGAACTTGGGCCGAACGTTGGCCAGTGTCTCTCTCCAGCCAGGCTGTGAGGTAGCT
GTTTCCAGGATGGGCACCTTTCCACACCCAGCAATGTGGCCAGGAGCCGCCATTCACGG
GTGCGACCAGCAGATGGCATCAGAGCCTCACTTTTGATGCACTCCGGCCACCAGCCACG
GGTCCAGGTTCTGGCCACCACCCAGGGTCTGAGCAGCTGCATCCTGCCCCTGCCGGGCA
CTCCCGGGGGCTGTGGGGCCTGTGGGGGCCCTGCCAGACACTCTTGGGGGCTGTGGGGG
GCCCTGCCAGGCACTCCCAGGGACTATGGGGGCTGTGGGGGGCCCTGCTGGGCACTCTG
AAGGGCATGGGGCTTAGGAATGAGAGGAGCTGTCTGATGATGATGGTGGGGGCACTGCA
GAGGCCCCCGGCCTGCTCAGGTCCAGTCTCGGCCCCTAAGTCAAGCCTCAGGCCAGCCT
CTCACCAGCCTGGGTTTCTCAGAGGGCCGGGACAAATGTTCTGGGTCTCTAATATTCCA
AGAAAGCCTCTGGCTGGACTCTGAGCCCCACCTGCGAG[ct]CCCTAGAATCACAGAGA
GCTAGGGTGAGAAGACCAGGGGGACTCCGTCCCACCCTCGTCGTGGCTGAGCCCACTGT
GGCCGGTGGTGGACCAGGCTGTGGCCTTTGCTGAGGGTCCCCAGGGCCCCTGGGGGCTA
CTGAGGCTGGAGGCCAGCGGTGGCCAGGAGGGTCCCTCCCTCAGCCACTCAAGCCAGAA
GGTCGAGTCCTGGTTTCTATGTGAGGAGGGGGCTTCAGGGGCTGGGACCTGGGGGCACC
GAAGGCCTGGAGCTGGGGTCCAGGCGGCTGAGGGTTAGTGCGTTCCCACGCTCCCCTCC
GCCAGCGCCGTGAGGAGAGGGAGGTCCACTCTGGAAAGAATGTTTGAGGGCAGGGGTAG
ACAGGGTCTGGGAACGCGGAG

agagactgaggccaggacaggcagtgtttaagcctgagtttctgggaaaggtagccctg
ggcagaacttgggccgaacgttggccagtgtctctctccagccaggctgtgaggtagct
gtttccaggatgggcacctttccacacccagcaatgtggccaggagccgccattcacgg
gtgcgaccagcagatggcatcagagcctcacttttgatgcactccggccaccagccacg
ggtccaggttctggccaccacccagggtctgagcagctgcatcctgcccctgccgggca
ctcccgggggctgtggggcctgtgggggccctgccagacactcttgggggctgtggggg
gccctgccaggcactcccagggactatgggggctgtggggggccctgctgggcactctg
aagggcatggggcttaggaatgagaggagctgtctgatgatgatggtgggggcactgca
gaggcccccggcctgctcaggtccagtctcggcccctaagtcaagcctcaggccagcct
ctcaccagcctgggtttctcagagggccgggacaaatgttctgggtctctaatattcca
aGAAAGCCTCTGGCTGGACTCTGAGCCCCACCTGCGAG[ct]CCCTAGAATCACAGAGA
GCTAGGGTGAGAAGACCAGGgggactccgtcccaccctcgtcgtggctgagcccactgt
ggccggtggtggaccaggctgtggcctttgctgagggtccccagggcccctgggggcta
ctgaggctggaggccagcggtggccaggagggtccctccctcagccactcaagccagaa
ggtcgagtcctggtttctatgtgaggagggggcttcaggggctgggacctgggggcacc
gaaggcctggagctggggtccaggcggctgagggttagtgcgttcccacgctcccctcc
gccagcgccgtgaggagagggaggtccactctggaaagaatgtttgagggcaggggtag
acagggtctgggaacgcggag

FIGURE 23hhhhh

67309

ATGCCCCTCCTAACATGAAAGGGATTTAAGCAAGCCAATTGCTTATTTCTGCCTGGGCC
AGGGACCCCAGTTCCTGACCTTCTCAAGAGATATGAACCTGACCCTTCTGAGTGTAGAA
CTGGGCTGTGGGGCCAGGAGATGTGGGTTTCAATCCCAGGACCCCCACTGGTGGCTGTG
CCATCTTGAGCAAGGCACTTTGTTTCTCCGAGTCTCTATTTCTTCACTGGTAAACAAAG
GCACAAATACCTCTTCACCACATCATAAGGGGATTAAATGATGTAGGAAAAAGGATGTT
GTATAGTCGTGCACATAGTAGGGCAGCAGGTCCAGGAGGTGGACGGCCCATCCAGGGAC
CCAGCGGAGCAGCCACTTCCCCACTTCTCAAGGGTGGTCACCAGGTATGTCCGCAGGGC
TGCCCCCTGCCCATCTCCAAGGCCTGACTGGCTGATCTCAGCTACACATTGGATACTAA
GTCCTAGGGCCAGAGCCAGCAGAGAGGTTTGCCTTACCTTGGAAGTGGACGTAGGTGTT
GAAAGCCAGGGTGCTGTCCACACTGGCTCCC[ga]TCAGGGAGCAGCCAGTCTTCCATC
CTGTCACAGCCTGCATGAACCTGTCAATCTTCTCAGCAGCAACATCCAGTTCTGTGAAG
TCCAGAGAGCGTGGGAGGACCACAGGGGTATAGAGAGCCAGGCCCTGCACAAACGGCTG
CTTCAGGTGCAGGCCTGGGGCTGTGAACACGCCCACCACCGTGGACAGCAGCAGCTGGG
CCTGGCTATCAGCCCTGCCCTGGGCCACTAGCAGGCCCTGTACAGCCTGCAGGGCAGAC
AGGACCTTGTGCGCATCCAGCCGGGAGGTGCAGTTCTTGTCCTTCCAAGGAACACCCAG
GATTGCCTGTAGCCTGTCAGCTGTGTGGTCCAAGGCTCCCAGATAGAGAGAGGCCAGGG
TGCCAAAGACAGCCGTTGGGGAGAGGACGGTGGCCCCATGGACCACGCCCCATAGCTCA
CTGTGCATGCCATATATACGG

atgcccctcctaacatgaaagggatttaagcaagccaattgcttatttctgcctgggcc
agggaccccagttcctgaccttctcaagagatatgaacctgacccttctgagtgtagaa
ctgggctgtggggccaggagatgtgggtttcaatcccaggacccccactggtggctgtg
ccatcttgagcaaggcactttgtttctccgagtctctatttcttcactggtaaacaaag
gcacaaatacctcttcaccacatcataaggggattaaatgatgtaggaaaaaggatgtt
gtatagtcgtgcacatagtagggcagcaggtccaggaggtggacggcccatccagggac
ccagcggagcagccacttccccacttctcaagggtggtcaccaggtatgtccgcagggc
tgccccctgcccatctccaaggcctgactggctgatctcagctacacattggatactaa
gtcctagggccagagccagcagagaggtttgccttaccttggaagtggacgtaggtgtt
gaAAGCCAGGGTGCTGTCCACACTGGCTCCC[ga]TCAGGGAGCAGCCAGTCTTCCATC
CTGTCacagcctgcatgaacctgtcaatcttctcagcagcaacatccagttctgtgaag
tccagagagcgtgggaggaccacaggggtatagagagccaggccctgcacaaacggctg
cttcaggtgcaggcctggggctgtgaacacgcccaccaccgtggacagcagcagctggg
cctggctatcagccctgccctgggccactagcaggccctgtacagcctgcagggcagac
aggaccttgtgcgcatccagccgggaggtgcagttcttgtccttccaaggaacacccag
gattgcctgtagcctgtcagctgtgtggtccaaggctcccagatagagagaggccaggg
tgccaaagacagccgttggggagaggacggtggccccatggaccacgccccatagctca
ctgtgcatgccatatatacgg

FIGURE 23iiiii

67320
AGGAGAGGAAGGGCGTGGAAACTGGAATGATCCTAGTGGGGTGTCTTGGCATCTCTTGG
CCTCATTTTCCCCATCTGAACCATGAAGCTAAAACTAGGGGATGTGGATTAAATGGTTC
CTACAACTACTTGCAAGGAGACCACTCTGTGTGGTTGCAAAGAACACTTTGAGAAGCTG
TGTGGGAAAGTTTCCTTCCTAGCAGGGTAGACTCAGCTAACTGCAGGTCATGTGGCCAT
TGTGGATGGGTTGGGAGCTCAAGTTTGGGGCAGAAGGGAATTTTTTTGGCAGCAGAGT
GGCAAGCCCTGCCGCCAGGCAAACTCTGCTCTTCCTCATCCTCAGAAGCACTTGCTCAC
TCTGCTAAATCAAAGTGAAACGCATGTTTACAGAATATTGGTCCAAAAGGGTCTCAGCA
TCTCCCACTACCCAGGGTGGCAGAGCCTCGGGCCGGCCTTGCTCCCCAAGAAGGGCTGA
CTGGGGCTCTGTCCCCTGCCCCAGGGCTCGAGGTAGTGTTTACAGCCCTCATGAACAGC
AAAGGCGTGAGCCTCTTCG [ag] CATCATCAACCCTGAGATTATCACTCGAGATGTGAG
TACAAAGCCCCCCTCACCAGCCCCTGTTCCTGGGGAGAGAGGCCCAGACAGGATTCCTG
GGGTGACTGGGGGCTGTTGGGGAGACAGACAGAGGGGCCTCTACCAGCTTGGCTCCCTC
CTGGTGGCCTGGGAGTCAGCCCAGCTCGCCCCTCTCTCCTACTGCCCCTCCCTTCAGGG
CTTCCTGCTGCTGCAGATGGACTTTGGCTTCCCTGAGCACCTGCTGGTGGATTTCCTCC
AGAGCTTGAGCTAGAAGTCTCCAAGGAGGTCGGGATGGGGCTTGTAGCAGAAGGCAAGC
ACCAGGCTCACAGCTGGAACCCTGGTGTCTCCTCCAGCGATGGTGGAAGTTGGGTTAGG
AGTACGGAGATGGAGATTGGCTCCCAACTCCTCCCTATCCTAAAGGCCCACTGGCATTA
AAGTGCTGTATCCAAGAGCTG

aggagaggaagggcgtggaaactggaatgatcctagtggggtgtcttggcatctcttgg
cctcattttccccatctgaaccatgaagctaaaactaggggatgtggattaaatggttc
ctacaactacttgcaaggagaccactctgtgtggttgcaaagaacactttgagaagctg
tgtgggaaagtttccttcctagcagggtagactcagctaactgcaggtcatgtggccat
tgtggatgggttgggagctcaagtttggggcagaagggaatttttttggcagcagagt
ggcaagccctgccgccaggcaaactctgctcttcctcatcctcagaagcacttgctcac
tctgctaaatcaaagtgaaacgcatgtttacagaatattggtccaaaagggtctcagca
tctcccactacccagggtggcagagcctcgggccggccttgctccccaagaagggctga
ctggggctctgtcccctgccccagggctcgaggTAGTGTTTACAGCCCTCATGAACAGC
AAAGGCGTGAGCCTCTTCG [ag] CATCATCAACCCTGAGATTATCACTCGAGATGTGAG
TACAAAGCCcccctcaccagcccctgttcctggggagagaggcccagacaggattcctg
gggtgactgggggctgttggggagacagacagaggggcctctaccagcttggctccctc
ctggtggcctgggagtcagcccagctcgcccctctctcctactgcccctcccttcaggg
cttcctgctgctgcagatggactttggcttccctgagcacctgctggtggatttcctcc
agagcttgagctagaagtctccaaggaggtcgggatggggcttgtagcagaaggcaagc
accaggctcacagctggaaccctggtgtctcctccagcgatggtggaagttgggttagg
agtacggagatggagattggctcccaactcctccctatcctaaaggcccactggcatta
aagtgctgtatccaagagctg

FIGURE 23jjjjj

67321
TTGGATAGACTGGGGGAAATAAGTCCTGTGGGACCTCCTGCCTTAAAGAAAGCAGGCGG
AGGGCCCTAAAGGAAATCAGGCAACCAGACCAAAAGAATGTGGACCAGGTGGTCCATGC
TGTGTCTCTTGTGACCCTTCTTCTCCCTGCCATGTCTTTTGGGAGAGCCCTTGTGTTGC
AAAAATGAGAGTGTGGTGGTATGGATTGGGGTTTAGGCAGAACAGTACTGGCCAAGCAG
CGCCTCCCTGGACCTCAATTTTCCCTCTGTGGAATGGGCTAGCAATCCTGGGCCTCCCC
AGGGCGAAGGAAAGACCACTCAGGAAGGGCACCGTCTGGGGCAGGAAAACGGAGTGGGT
TGGATGTATTTTTTTCACGGATGGGCATGAGGATGAATGCTTGTCCAGGCCGTGCAGCA
TCTGCCTTGTGGGTCACTTCTGTGCTCCAGGGAGGACTCACCATGGGCATTTGATTGGC
AGAGCAGCTCCGAGTCCGTCCAGAGCTTCCTGCAGTCAATGATCACCGCTGTGGGCATC
CCTGAGGTCATGTCTC [ga] TAAGTGTGGGCTGGAGGGGAAACTGGGTGCCGAGGCTGA
CAGAGCTTCCCATTTCACCTTGTGGGCCCTTCCCAGGCAGAGCTTCAGGTGCCCCTCTT
CCCAGTCATTGATACTTAGCGGTCCTGGCCCCCTTTCCTCTCCCTGCTGGTGGTATTGC
ACGCCAATGACTCGGCCAGATGCCCAGACCCCTGTTCTTGGTTTACCTGCAGAATATTA
TCTTTGCCACCCCGCGGGATGGCTCAACCCACTTTCAGGATGCAGGTCTCCTAATAGCA
ACCTGATATAGCAGAAAGACCCCTGGGCTGGGAGTCTGAGACCTAGTTCTAGCCCAGCC
CTGAACCTCAGTTTCCCTTTCTGTGAAACAAGAATGTTGAACTTGATGATTCCCAATTT
TCCTTTTGACCTTGAAATGGTAGAATATTTATCCCTTTGAGGTGACTCGGATGGTAGAC
TCTCAGACACCATAGCACACG

ttggatagactgggggaaataagtcctgtgggacctcctgccttaaagaaagcaggcgg
agggccctaaaggaaatcaggcaaccagaccaaaagaatgtggaccaggtggtccatgc
tgtgtctcttgtgacccttcttctccctgccatgtcttttgggagagcccttgtgttgc
aaaaatgagagtgtggtggtatggattggggtttaggcagaacagtactggccaagcag
cgcctccctggacctcaattttccctctgtggaatgggctagcaatcctgggcctcccc
agggcgaaggaaagaccactcaggaagggcaccgtctggggcaggaaaacggagtgggt
tggatgtatttttttcacggatgggcatgaggatgaatgcttgtccaggccgtgcagca
tctgccttgtgggtcacttctgtgctccagggaggactcaccatgggcatttgattggc
agagcagctccgagtccgtccagagcttcctgcagtcaatgatcacCGCTGTGGGCATC
CCTGAGGTCATGTCTC [ga] TAAGTGTGGGCTGGAGGGGAAACTGGGTGccgaggctga
cagagcttcccatttcaccttgtgggcccttcccaggcagagcttcaggtgcccctctt
cccagtcattgatacttagcggtcctggccccctttcctctccctgctggtggtattgc
acgccaatgactcggccagatgcccagacccctgttcttggtttacctgcagaatatta
tctttgccaccccgcgggatggctcaacccactttcaggatgcaggtctcctaatagca
acctgatatagcagaaagacccctgggctgggagtctgagacctagttctagcccagcc
ctgaacctcagtttccctttctgtgaaacaagaatgttgaacttgatgattcccaattt
tccttttgaccttgaaatggtagaatatttatccctttgaggtgactcggatggtagac
tctcagacaccatagcacacg

FIGURE 23kkkkk

67325
GGGGCAGGGCTGGTGGTCAGCTGGGGCGGGGTGGGAGCTGGAGGTCCGTGGTCACCAGC
TGCCCTGACTAATGTCGTTACTTGAATATAACCCTGTGAAGGCAGGAACCACGTCTGTC
TGGTTCACTTCCCACGGTGGTTGAGACATAGTGGGCACTCCGGAAGTATTTGTTGAATG
AGTGAAAGCCCCGCTGGGGGAAACTGGGTACAGCTCTTTCCTCAGTTTCCCCATCTGCA
CTCTGGGCTGAATGCTGGGGCTCCTCCCAATCTCCCTGAAGCTGGACCTGAGCCCAGTA
GGGACACACAGGGTCCAGCCAGCGTCCTGGCTTCCTCCAGGGTCATTTCATCTACAAGA
ATGTCTCAGAGGACCTCCCCCTCCCCACCTTCTCGCCCACACTGCTGGGGGACTCCCGC
ATGCTGTACTTCTGGTTCTCTGAGCGAGTCTTCCACTCGCTGGCCAAGGTAGCTTTCCA
GGATGGCCGCCTCATGCTCAGCCTGATGGGAGACGAGTTCAAGGTGAGTGGGTGGGGCT
GGGCTGCTAGGG[ga]ATCCAGATGGCATGTGGTATGTGTGTGTGTGCACACGCATGGG
GAGGAGGGAGGAAACTCGGAAACTTGGTGGTGGGCAAAAGAACTAAGCTGGAGCAATAG
CAGTGAAGTCCAGACTGGGCACAGTGGCTCACACCTGTAATCCCAATCCTTTGGGAGGC
TGAGATGTAGCAGGACGAACCGCAGACAAAACTCCTCAGACACTGAGTTAAAGAAGGAA
AGAGTTTATTCAGCCGGGAGCATGGGTAAGACTCCTGTCTCAAGAGCGGAGCTCTCCGA
GTGAGCAATTCCTGTCCCTTTTAAGGGCTCACAACTCTAAGGGGGTCTGCATGAGAGGG
TCGTGATCTATTGAGCAAGTAGCAGGTACGTGACTGGGGGCTGCATGCACCGGTAATCA
GAACGAAACAGAACAGGACAGGGATTTTTACAATGCTCTTTCATGCAATGTCTGGAATC
TATAGATAACATAACTGGTTA

ggggcagggctggtggtcagctggggcggggtgggagctggaggtccgtggtcaccagc
tgccctgactaatgtcgttacttgaatataaccctgtgaaggcaggaaccacgtctgtc
tggttcacttcccacggtggttgagacatagtgggcactccggaagtatttgttgaatg
agtgaaagccccgctgggggaaactgggtacagctctttcctcagtttccccatctgca
ctctgggctgaatgctggggctcctcccaatctccctgaagctggacctgagcccagta
gggacacacagggtccagccagcgtcctggcttcctccagggtcatttcatctacaaga
atgtctcagaggacctccccctccccaccttctcgcccacactgctgggggactcccgc
atgctgtacttctggttctctgagcgagtcttccactcgctggccaaggtagctttcca
ggatggccgcctcatgctcagcctgatgggagacgagtTCAAGGTGAGTGGGTGGGGCT
GGGCTGCTAGGG[ga]ATCCAGATGGCATGTGGTATGTGTGTGTGTGCAcacgcatggg
gaggagggaggaaactcggaaacttggtggtgggcaaaagaactaagctggagcaatag
cagtgaagtccagactgggcacagtggctcacacctgtaatcccaatcctttgggaggc
tgagatgtagcaggacgaaccgcagacaaaactcctcagacactgagttaaagaaggaa
agagtttattcagccgggagcatgggtaagactcctgtctcaagagcggagctctccga
gtgagcaattcctgtcccttttaagggctcacaactctaagggggtctgcatgagaggg
tcgtgatctattgagcaagtagcaggtacgtgactgggggctgcatgcaccggtaatca
gaacgaaacagaacaggacagggatttttacaatgctctttcatgcaatgtctggaatc
tatagataacataactggtta

FIGURE 2311111

67329
GCAAATCCATAGAGACAGAAAGCACATTTATGGTTGCCAGGAGCTGGGAAAGGGCAGGA
TGGGGAATGACTGTTTATTGGATGTGGGGCTCTATTTTGGGGTGATGAGAATGTTCTGG
AATTAAATTCATGGCTGCATAACACTGTGAACATACTAAATGCCCCTGAATTGTACACT
TTAAAATGGTTAAAGTGGCAAGTTTTCACTAAGCAGTAAATTAAATTCTACTACAATTT
TAAAAAGACTAAAAAATAATTTAAAAAAGATTAAATGAGATAACGCAAAAAAGCATTAT
CTCGAAAATACAGCTGATATTAGTATAATTCTTACTAAGTTTTAAGAGTCTAAGGTGCA
GGATTCTAAGTTTAAAGGGATAGGCTCTTTTGGTTTTTTGGTTTAGTTATTTGGTTTTT
TTTTTTAATCCATTATCCCCACCCTTGGGAGGCCCCCAGCACCCAGTCTGCACTAGAGG
ATGGGGCCCACCTCCCTTTTCTCTCCAGGCCCAGCCACTGACCACCAGTACCCTGGCCA
GGGGCACCCT[tc]GGTCATTGCCCTCCGTGGCCCAAGGAAGGGAACAGAAACAACAGC
CAAGAAGACAATAGCCGCCGGGAAGTCCTCACATTTCTGGAGAAATAGAGCCCATTAAT
GAATGAAGTTCCTCCAGCCTGATCGGAGGACGGGGTGCTGGGGAGGCCTGGGCTAAAGG
GCTCACCTCCAGCCCCCACCCTGGCAGGGCCGATGGTACATGCTCACTCAGTGAGGGGG
CTCCAGAGGTCTGTGGGTACGAACCCAAGGGCTGGTGCCCAGGGGCAATCAGCTTATGT
CTCTGAGCCTTGGGAAACAGTGAGGGTCAGCCCGGCTCCCCACGTGCTTCTGGGCAGCT
TTGGTATTGGAGCAGGTGCAAACTCGGGACTAGGGCAGGACCCCCTGAGAGGCGACTGA
GCAAGGCCATCCCGACTCATGTTTCCTTGGCCCTGCCCGGGGCACAGCATCCTGCCCAC
ATCCCTGCAGCCCTGGCTCCT

gcaaatccatagagacagaaagcacatttatggttgccaggagctgggaaagggcagga
tggggaatgactgtttattggatgtggggctctattttggggtgatgagaatgttctgg
aattaaattcatggctgcataacactgtgaacatactaaatgcccctgaattgtacact
ttaaaatggttaaagtggcaagttttcactaagcagtaaattaaattctactacaattt
taaaaagactaaaaaataatttaaaaaagattaaatgagataacgcaaaaaagcattat
ctcgaaaatacagctgatattagtataattcttactaagttttaagagtctaaggtgca
ggattctaagtttaaagggataggctcttttggttttttggtttagttatttggttttt
ttttttaatccattatccccacccttgggaggcccccagcacccagtctgcactagagg
atggggcccacctcccttttctctccaggcccagccactgaccaCCAGTACCCTGGCCA
GGGGCACCCT[tc]GGTCATTGCCCTCCGTGGCCCAAGGaagggaacagaaacaacagc
caagaagacaatagccgccgggaagtcctcacatttctggagaaatagagcccattaat
gaatgaagttcctccagcctgatcggaggacggggtgctggggaggcctgggctaaagg
gctcacctccagcccccaccctggcagggccgatggtacatgctcactcagtgaggggg
ctccagaggtctgtgggtacgaacccaagggctggtgcccaggggcaatcagcttatgt
ctctgagccttgggaaacagtgagggtcagcccggctccccacgtgcttctgggcagct
ttggtattggagcaggtgcaaactcgggactagggcaggaccccctgagaggcgactga
gcaaggccatcccgactcatgtttccttggccctgcccggggcacagcatcctgcccac
atccctgcagccctggctcct

FIGURE 23mmmmm

67330
GGGAACTAGTGCCGCCCCAGGGCCCCAAGGTGGGCGGTTCGGTGATTCAGAGAGGGCAG
CTCTGTGTTAGGACACACTGGGGCCAGCCAGGAAGGGTGGAAAAGATAGGGACCAGCGT
GAGCATAGAGGCTAAGGGACCATGGGAGCTCCAAGCGCGCTCACAGTGGGGACCAGGTC
CTGGGGGCTGGGGACACCAGGGAGGTGAAATACCCCTCCAGCGGGTAGGGAGGGTGGGC
AGAGGAGGGCCAGCGGCCAGGCATTTGGGAGGGGCTCCTGCTCTTTGGGAGAGGTGGGG
GGCCGTGCCTGGGGATCCAAGTTCCCCTCTCTCCACCTGTGCTCACCTCTCCTCCGTCC
CCAACCCTGCACAGGCAAGATCGTGGACGCCGTGATTCAGGAGCACCAGCCCTCCGTGC
TGCTGGAGCTGGGGGCCTACTGTGGCTACTCAGCTGTGCGCATGGCCCGCCTGCTGTCA
CCAGGGGCGAGGCTCATCACCATCGAGATCAACCCCGACTGTGCCGCCATCACCCAGCG
GATGGTGGATTTCGCTGGC [ag] TGAAGGACAAGGTGTGCATGCCTGACCCGTTGTCAG
ACCTGGAAAAAGGGCCGGCTGTGGGCAGGGAGGGCATGCGCACTTTGTCCTCCCCACCA
GGTGTTCACACCACGTTCACTGAAAACCCACTATCACCAGGCCCCTCAGTGCTTCCCAG
CCTGGGGCTGAGGAAAGACCCCCCCAGCAGCTCAGTGAGGGTCTCACAGCTCTGGGTAA
ACTGCCAAGGTGGCACCAGGAGGGGCAGGGACAGAGTGGGGCCTTGTCATCCCAGAACC
CTAAAGAAAACTGATGAATGCTTGTATGGGTGTGTAAAGATGGCCTCCTGTCTGTGTGG
GCGTGGGCACTGACAGGCGCTGTTGTATAGGTGTGTAGGGATGGCCTCCTGTCTGTGAG
GACGTGGGCACTGACAGGCGCTGTTCCAGGTCACCCTTGTGGTTGGAGCGTCCCAGGAC
ATCATCCCCCAGCTGAAGAAG

gggaactagtgccgccccagggccccaaggtgggcggttcggtgattcagagagggcag
ctctgtgttaggacacactggggccagccaggaagggtggaaaagatagggaccagcgt
gagcatagaggctaagggaccatgggagctccaagcgcgctcacagtggggaccaggtc
ctgggggctggggacaccagggaggtgaaatacccctccagcgggtagggagggtgggc
agaggagggccagcggccaggcatttgggaggggctcctgctctttgggagaggtgggg
ggccgtgcctggggatccaagttcccctctctccacctgtgctcacctctcctccgtcc
ccaaccctgcacaggcaagatcgtggacgccgtgattcaggagcaccagccctccgtgc
tgctggagctgggggcctactgtggctactcagctgtgcgcatggcccgcctgctgtca
ccaggggcgaggctcatcaccatcgagatcaaccccgactgtgccgccatCACCCAGCG
GATGGTGGATTTCGCTGGC [ag] TGAAGGACAAGGTGTGCATGCCTGACCCgttgtcag
acctggaaaaagggccggctgtgggcagggagggcatgcgcactttgtcctccccacca
ggtgttcacaccacgttcactgaaaacccactatcaccaggcccctcagtgcttcccag
cctggggctgaggaaagacccccccagcagctcagtgagggtctcacagctctgggtaa
actgccaaggtggcaccaggaggggcagggacagagtggggccttgtcatcccagaacc
ctaaagaaaactgatgaatgcttgtatgggtgtgtaaagatggcctcctgtctgtgtgg
gcgtgggcactgacaggcgctgttgtataggtgtgtagggatggcctcctgtctgtgag
gacgtgggcactgacaggcgctgttccaggtcacccttgtggttggagcgtcccaggac
atcatcccccagctgaagaag

FIGURE 23nnnnn

67340
AGCTTCCTGAGTAGCTGGGATTACAGGCACTCACCTCCACGCCCAGCTAACTTTTGTAT
TTTTAGTACAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCG
TGATCCGCCCTCCTCGCCTCCCAAAGTGCTGTGATTACAGGAGTGAGCCACCGCTCCTG
GCCAGAAATCTCTTCTTTATTATGTCTACTGTCCGTTATCCAACTCCAGAAGGTAAGAA
CCTCCACTGATACATAAGGACTTGTATACCCCACGTGCCTGCAACAGTGCTTGGCACCT
AGTAGGCATACCAAAATATATAAATGTTGAACAAATGAAGAAAGTTAAAGTAAAACTAG
AGGTCCAAAAATATCACAAAAGCCATCTATGGTCGCCTTTTCCCTACCTGATTTTGCTG
AGTGGCCTTACTTTTCAGTCCTCTACACAGCTGGAACATTAATGAACACAGAGGGGGAA
GAAGTGTGTTTACTCTAGGATCACCTCTCAATGGGTCACTTGGCAAGGGCATCTTTGCT
TCTTCGTCAGCTCCTTTTGACA[ct]GGGGGTGAAGGGTTTTCTGCACCACACTTTGAC
CACAAGCATCACCAATTTCACTGAACCCAACAGAAATTTGGACCCTCTGGGGGCTCTCT
GCGTGGCAGGGCCCTTTTCTTTTTCTTTGGGCTTAGGCTGCAATTTGAAACACCACTTT
CCTGAGCCAGCATCCCCCTTGCAGCGCTGTCACAGGGAGGCTTAGGCAGCCACGTGGAA
GCCACCTACCCCGACCTTTGGCAGAATTTCCAAACACAACACAGTAGCTTTAAGTTGAT
TAATTTGGAACTCTGACCTTGGCCCCAAAAGGTAAGAATACATAACAAGGTATTTTATT
CTCAAAATGTGTCAGGATAAGAAGCACTTCTGTAAATCGACCTTTTTAAAATAGATATA
ATTAGATTTGCAGTTGGGGGCAGTAAAGAAAGGGTCTGAACAGTGGATAACATGTTGAG
AGGTTAATTATTAATGGGCAG

agcttcctgagtagctgggattacaggcactcacctccacgcccagctaacttttgtat
ttttagtacagatggggtttcaccatgttggtcaggctggtctcgaactcctgacctcg
tgatccgccctcctcgcctcccaaagtgctgtgattacaggagtgagccaccgctcctg
gccagaaatctcttctttattatgtctactgtccgttatccaactccagaaggtaagaa
cctccactgatacataaggacttgtataccccacgtgcctgcaacagtgcttggcacct
agtaggcataccaaaatatataaatgttgaacaaatgaagaaagttaaagtaaaactag
aggtccaaaaatatcacaaaagccatctatggtcgccttttccctacctgattttgctg
agtggccttacttttcagtcctctacacagctggaacattaatgaacacagagggggaa
gaagtgtgtttactctaggatcacctctcaatgggtcacttggcaagggcATCTTTGCT
TCTTCGTCAGCTCCTTTTGACA[ct]GGGGGTGAAGGGTTTTCTGCACCACACTTTGac
cacaagcatcaccaatttcactgaacccaacagaaatttggaccctctgggggctctct
gcgtggcagggcccttttctttttctttgggcttaggctgcaatttgaaacaccacttt
cctgagccagcatcccccttgcagcgctgtcacagggaggcttaggcagccacgtggaa
gccacctaccccgacctttggcagaatttccaaacacaacacagtagctttaagttgat
taatttggaactctgaccttggccccaaaaggtaagaatacataacaaggtattttatt
ctcaaaatgtgtcaggataagaagcacttctgtaaatcgacctttttaaaatagatata
attagatttgcagttgggggcagtaaagaaagggtctgaacagtggataacatgttgag
aggttaattattaatgggcag

FIGURE 2300000

67342
GCAGCCTGTTGTGCCTTGTGCCTCGAAGAGGTTTGGTATCTGCCAGTTTCTCCCTCGCT
GTTTTTATGGCTTTCAAAAGCAGAAGTAGGAGGCTGAGAAATTTCTCTGTTGAATACCT
GATTTCACAATCAAGTTAAAGGAAAGGGGAAAAGAGTATTGGTGGAAGCTTCTTAGGGG
AGGGGACTAATAAACTGAGATAATTCTCTGGTTCATGGAAGGGCAAGGAGTAGCAAACT
ATGACACATTTTGCAAATGTATCACCATGCAAATATGCATTGTTTTCCTGACAATCGTT
GTGCAGTTGATGTCCACATTAAAATACTGGATTTTCCCACGTTAGAAGAATGTTTAAAT
TTAGTATATGTGGGACAAAGTGGAAGACACACAGATTTATACATGCACATACTTTTCTT
CATTCACTTCTTTGTACTTAAGTTTAGGAATCTTCCCACTTACAGATGGATAAATGGGT
ACAATGAAGGGCCAATAGCCCTCCCTGTCTGTATTGAGGGTGTGGGTCTCTACCTTGGG
TGCTGTTCTCTGCCTC [ag] GGAGCTCTCTGTCAATTGCAGGAGCCTCTGAGGAGAAAA
TTGACCTTTCTTGGCTGGGGCAGAGAACATACGGTATGCAGGGTTCAGGCTCCTGACGG
AGTTGGGGCAACCCTGGAGATAAGCTCACACAACCCTGCAAGACCAGGTGCTGTTACCC
TAGCCAATCTCATGGATGAACCAGATCAATGCCAGATGAGCTCTGCCTAAAATGATTTT
TTGGTGAACTCTGAAAAGTGGAATATTGTTTCTGTAAGAATATCCATCTGAGACTCTAT
CTCTTGGTAATACCAACCAAGAGTTATCAGTTTCTCTTTAACCGAGACACCAGCAAAGT
GCCTGCTCCAGGGTACTGCCCAGGGGAGCCCTCCATTTGTAGAATGAATGAGAGTCCAG
GTTATGAACAGTGCCTGGAGTGTAGGAACACCCTCCTTTGCCTCTTTGACAGGTCTGCA
TCATAACACTTTTTTTTTTTT

gcagcctgttgtgccttgtgcctcgaagaggtttggtatctgccagtttctccctcgct
gtttttatggctttcaaaagcagaagtaggaggctgagaaatttctctgttgaatacct
gatttcacaatcaagttaaaggaaaggggaaaagagtattggtggaagcttcttagggg
aggggactaataaactgagataattctctggttcatggaagggcaaggagtagcaaact
atgacacattttgcaaatgtatcaccatgcaaatatgcattgttttcctgacaatcgtt
gtgcagttgatgtccacattaaaatactggattttcccacgttagaagaatgtttaaat
ttagtatatgtgggacaaagtggaagacacacagatttatacatgcacatacttttctt
cattcacttctttgtacttaagtttaggaatcttcccacttacagatggataaatgggt
acaatgaagggccaatagccctccctgtctgtattgagggtgtGGGTCTCTACCTTGGG
TGCTGTTCTCTGCCTC [ag] GGAGCTCTCTGTCAATTGCAGGAGCCTCTGAGgagaaaa
ttgacctttcttggctggggcagagaacatacggtatgcagggttcaggctcctgacgg
agttggggcaaccctggagataagctcacacaaccctgcaagaccaggtgctgttaccc
tagccaatctcatggatgaaccagatcaatgccagatgagctctgcctaaaatgatttt
ttggtgaactctgaaaagtggaatattgtttctgtaagaatatccatctgagactctat
ctcttggtaataccaaccaagagttatcagtttctctttaaccgagacaccagcaaagt
gcctgctccagggtactgcccaggggagccctccatttgtagaatgaatgagagtccag
gttatgaacagtgcctggagtgtaggaacaccctcctttgcctctttgacaggtctgca
tcataacacttttttttttttt

FIGURE 23ppppp

67356
GAAATACCATATTGCATCAAACCTAAGACGCCATCAAGAATAAAAGGCACTTTTCTTTA
CATTACTACCCAGACGCAAACAGAGCTGCCAATTCAACCATGATGAGTCACCAGTTATA
GGAGGTTTGATTTCAGAGCTATAAGAGTGTATGTCCTAGAACCAATGAGCTATCGTAGA
TCCAAGAATCTACATATCTGAGTTGGAAGGGCTGCCAGCCCTTGGGGCATGATCTTCCA
TCCTCAAAGACTTCTTCAGATTTGAAGAGCAAGGGGAAGGACTGCCTGGTGTCTTAACG
AAGTGTCTCCTACTCAGCCAGTAGGACCCTGAGCACTCTGGGGCATCCTGGCATCTGTT
GCCCAGCTAATGGTTCCCACCAGTCACCCGTCCCAACCCATGCCACCATCCAGTGCCCA
GCAGCTCTCAGAGATACTCACTTACTACAGGAGACACACTCGTTTTCTCTTAGAAAGAA
ACCTGCATGGCAGGTGCACACGGTGTTCTGTTTCTCCTGGCCTGTAGGGAGAAGTGCGG
CACAGCTAAAGGAG[ag]AGCGCCTGCACCCCCACCCCACAGGACAGAGGAAGTGACGA
GGGACAGGGTGGGGGCGGCCAGAGAGGAGTTGGTTGTCAGACCCACAGAATACAGGAGG
GGGAAGGAAAGGAAGTGCCACCGCATGGGGAAGGGGCCAACCCCTGGGGTGGGGAGAGG
GCTTGGCCTCAGGAGAGCTGCGCTCACAGGAGAGGTGCACGGTCCCATTGAGGCAGAGG
CTGCAATTGAAGCACTGGAAAAGGTTTTCACTCCAATAATGCCGGTACTGGTTCTTCCT
GCAGCCACACACGGTGTCCCGGTCCACTGTGCAAGAAGAGATCTCCACCTGACCCATTT
CTGGTGAGGGGAGAAGATGGGGTATGAGTCCTGCATCCTCCTGTCCCTGCATCCCCTTC
CTGACATACCCCTAAGTGTGTGTCTCTGTAATACACACTCACATCCATGCAGTGTCCCA
CCAAAACACACACCTTCCTGC

gaaataccatattgcatcaaacctaagacgccatcaagaataaaaggcacttttcttta
cattactacccagacgcaaacagagctgccaattcaaccatgatgagtcaccagttata
ggaggtttgatttcagagctataagagtgtatgtcctagaaccaatgagctatcgtaga
tccaagaatctacatatctgagttggaagggctgccagcccttggggcatgatcttcca
tcctcaaagacttcttcagatttgaagagcaaggggaaggactgcctggtgtcttaacg
aagtgtctcctactcagccagtaggaccctgagcactctggggcatcctggcatctgtt
gcccagctaatggttcccaccagtcacccgtcccaacccatgccaccatccagtgccca
gcagctctcagagatactcacttactacaggagacacactcgttttctcttagaaagaa
acctgcatggcaggtgcacacggtgttctgtttctcctggcctgtagggagaagTGCGG
CACAGCTAAAGGAG[ag]AGCGCCTGCACCCCCACCCcacaggacagaggaagtgacga
gggacagggtgggggcggccagagaggagttggttgtcagacccacagaatacaggagg
gggaaggaaaggaagtgccaccgcatggggaaggggccaacccctggggtggggagagg
gcttggcctcaggagagctgcgctcacaggagaggtgcacggtcccattgaggcagagg
ctgcaattgaagcactggaaaaggttttcactccaataatgccggtactggttcttcct
gcagccacacacggtgtcccggtccactgtgcaagaagagatctccacctgacccattt
ctggtgaggggagaagatggggtatgagtcctgcatcctcctgtccctgcatccccttc
ctgacatacccctaagtgtgtgtctctgtaatacacactcacatccatgcagtgtccca
ccaaaacacacaccttcctgc

FIGURE 23qqqqq

67363
AGATCCAGAAGCTGAAGGAACAGATGAGGAGGCACCAAGAGAGCCTTGGAGGAGGTGCC
TAAGTTTCCCCCAGTGCCCACAGCACCCTCCGGCACTGAAAATACACGCACCACCCACC
AGGAGCCTTGGGATCATAAACACCCCAGCGTCTTCCCAGGCCAGAGAAAGTGGAAGAGA
CCACAAACCGCAGGCAATTGGCAGGCAGTGGGGGAGCCAGGGCTCTGCAGTCTTAGTCC
CATTCCCCTTTGATCTCACAGCAGGCAGGGCACCCAGGCCTTATAGGAATTCACCCTGG
ACCATGCCCTAAAATAACCTCACCCCAAATACAATAAAGGGACGAAGCACTTATAGATA
CCACAGACACATGTGTTTCATTTTTAGTTTTGTTAAAAAAAAATTCTGACAAATCAGAA
ATGGGGGTTCAGGAGTGGTGGTGATGCAAAAGATGGAAGCCATGGGGTGGGGGCTGTCA
GGGGTGGGGGCAGTAGTGTCTCCTTCACCCCCACCCTGGTGTCCTCTCCTGAAGGACAG
ACGGTCACATTCCAAAATGGG[ca]GAGTCTTCTACCGTGTCTGTTCAACTGAGAAGAA
AACGTAGCATGGTCAGAATAAGGCATGAAAAGGGGAAAGTGAGGCAGGAACACACGGCA
CACATGCAGACACTGGTGTACTGCCTGGGTTCAGAGGACGGACGTGGGGGTGAGGGAAG
GGATGTAATATGATGAGAGAAGACAGAAACCCCACATAAAGGTCAGAAAAACATCCCAA
CACAGCATCAAAGACCAGGGGGCATGAACCAGTCAAGTGTCCATTATGCATCAGATGCC
CATGACCTATGTGATGGGATTTAGGACAAACACACTAAGGAACAGGGAGGACCTAAAGG
GTTTCATGAGATCAGTACTCACTGTAGGAGGAGATGTCTATCTCATCAGGCAGCTCACT
AATATTGACCTCAAAGCGATCCTGCACATCATTGAGGATCTTGGCATCATTCTCATCGG
ACACAAATGTGATAGCCAAGC

agatccagaagctgaaggaacagatgaggaggcaccaagagagccttggaggaggtgcc
taagtttcccccagtgcccacagcaccctccggcactgaaaatacacgcaccacccacc
aggagccttgggatcataaacaccccagcgtcttcccaggccagagaaagtggaagaga
ccacaaaccgcaggcaattggcaggcagtgggggagccagggctctgcagtcttagtcc
cattcccctttgatctcacagcaggcagggcacccaggccttataggaattcaccctgg
accatgccctaaaataacctcaccccaaatacaataaagggacgaagcacttatagata
ccacagacacatgtgtttcatttttagttttgttaaaaaaaaattctgacaaatcagaa
atgggggttcaggagtggtggtgatgcaaaagatggaagccatggggtgggggctgtca
ggggtgggggcagtagtgtctccttcacccccaccCTGGTGTCCTCTCCTGAAGGACAG
ACGGTCACATTCCAAAATGGG[ca]GAGTCTTCTACCGTGTCTGTTCAACTGAGAAGAA
AACGTAGCATGgtcagaataaggcatgaaaaggggaaagtgaggcaggaacacacggca
cacatgcagacactggtgtactgcctgggttcagaggacggacgtgggggtgagggaag
ggatgtaatatgatgagagaagacagaaaccccacataaaggtcagaaaaacatcccaa
cacagcatcaaagaccagggggcatgaaccagtcaagtgtccattatgcatcagatgcc
catgacctatgtgatgggatttaggacaaacacactaaggaacagggaggacctaaagg
gtttcatgagatcagtactcactgtaggaggagatgtctatctcatcaggcagctcact
aatattgacctcaaagcgatcctgcacatcattgaggatcttggcatcattctcatcgg
acacaaatgtgatagccaagc

FIGURE 23rrrrr

67364
AGGTGTGTGCCACCATGCACGGCTAATTTTTGCATTTTTAGTAGAGAGAGGGTTTCATC
CTGTTGGCCACATTGGTCTTAAACTCCTGACCTCAAATAATCCACACGCCTTGGCCTCC
CAAACTGCTGAGATTACAGGTGTAAGCCATTGTGCACTTGGCCAGAATCCTCAATATTC
ACACACCACTGGAGCTGTTTTAAAGTTTCCGGCTTTCTCTGCCACATACCCCAAAATTA
TTAAACTGATATGATTCAAAGTCAGTATAAAGTAGTAAGAAAAGGGTGGTCTTGTGTTA
AGCATCATCCATAGCCCAATTACGAATCCTCCTGTTACATAGGAACTCAACACTCTGTT
ACACCACAGCAAACTAAAGCTTCTCCAAAATTAAAGAGACTATTGGCCTACAAGTTTCT
TATCCCTCCAACTTGCCACACCCTCACTCTCAGGTCTCTTTACCTTGGCTTACCTTGAC
ATTGGGCATGTATTTAGAGAAGCGCTCATATTCCTTGCTGATCTGAAAAGCCAACTCCC
GAGTGTGACACATCACCAG[ct]ACAGACACCTTAGGCAGGAAGTATACGGAGACATAT
GGTAAATGTAGCTCTTCATTATCCCCTCTAGGGAAGTGACTGTCACAAAAACACACCTG
GGCCGATAATAAATGACTTCAATTCTGTGATCTAAATCATGAACCCCACGCTTGCGACA
GAACATCCCCCACAGCTGTCAGGTTGTCAAGGGTAACAGAGGTCATGTGCTCATGGCTC
TGCAAGCATCATGTAGTTAGGACAAAAACACCCTTCCCTTATAGTCCTAACCAAAATCC
CCTCCCCAGCACTCTCCCCAAATATACCTGCCCAGTAACTGGCTCCAGCTGTTGCAGTG
TGGCCAAGACAAACACTGCTGTCTTTCCCATGCCCGACTTGGCCTGGCACAGGACATCC
ATTCCCAGAATGGCCTGAGGGATGCACTCATGCTGGACTAAAAGTTGGGGGGGGAGGAA
GATAAATTAGACTTCAGTCTC

aggtgtgtgccaccatgcacggctaatttttgcatttttagtagagagagggtttcatc
ctgttggccacattggtcttaaactcctgacctcaaataatccacacgccttggcctcc
caaactgctgagattacaggtgtaagccattgtgcacttggccagaatcctcaatattc
acacaccactggagctgttttaaagtttccggctttctctgccacataccccaaaatta
ttaaactgatatgattcaaagtcagtataaagtagtaagaaaagggtggtcttgtgtta
agcatcatccatagcccaattacgaatcctcctgttacataggaactcaacactctgtt
acaccacagcaaactaaagcttctccaaaattaaagagactattggcctacaagtttct
tatccctccaacttgccacaccctcactctcaggtctctttaccttggcttaccttgac
attgggcatgtatttagagaagcgctcatattccttgctgatctgaAAAGCCAACTCCC
GAGTGTGACACATCACCAG[ct]ACAGACACCTTAGGCAGGAAGTATACGGAGACatat
ggtaaatgtagctcttcattatcccctctagggaagtgactgtcacaaaaacacacctg
ggccgataataaatgacttcaattctgtgatctaaatcatgaaccccacgcttgcgaca
gaacatcccccacagctgtcaggttgtcaagggtaacagaggtcatgtgctcatggctc
tgcaagcatcatgtagttaggacaaaaacacccttcccttatagtcctaaccaaaatcc
cctccccagcactctccccaaatatacctgcccagtaactggctccagctgttgcagtg
tggccaagacaaacactgctgtctttcccatgcccgacttggcctggcacaggacatcc
attcccagaatggcctgagggatgcactcatgctggactaaaagttggggggggaggaa
gataaattagacttcagtctc

FIGURE 23sssss

67371
GTGTAATGTATTAGAGCAAATCCTCTTGATTAGGCTTGAGAATGGAGCCATGGAGCCCC
ATTTTTTTCCCACCCTTCATGCAGTAGTGTTTAATTAAATATTTAAAATATTTAATGCC
CTGCACAGGCATCATTTAATTGGAATGAACAACTGCTAACTGCTGGCACAGGGCTCTAG
AAGGCCCCAGATATCAGTAATTTACCACTGTTTGCTTGCTCTTGGGATAGGAAGGATCC
GGGGATCCTAGAGGAGGAGCTAGGGCAGTTGGGTGCTGGAGGAGGCACATGGGGGCTCA
GCACAGCCACTTGTTTGCCAGCTGGTGGAGCAGTGTGGAACTCGCCTTCTTGGGAGGAA
GAAACACGTCTCCAGACTTCCATAACAAAGTACCCAGAGTTGCTGGGCTAGTTACAGTT
CCAATGACCATTCCTCCCCAGCAGGATAAGCCCAGGGCCCCACCCTACCTGGGTCCCCC
TTCTCGCCCCGAGGGCCCTCTCTCCCATCCCGTCCATCGCGACCAGGCAGGCCACTCTC
CACTGAGCTACACATGACCAGGGTGCAAGCACTGGGC[ga]TTGTTCTGTGGGAGTAGG
TCTTCATTTCTGCTTCCAGGTAGCCCAGGGGCTGTGTGAGCAGGACCAGTGCAGAGAGG
AGGAAGAGCAGCATGGCCTGGAGAGGTGAACAGAAAGAGAAAAGACATGCTTATGCTTC
ATGGACATGGTTTAGGGCTTGGCTCAGCTTCTAGAGGTGACAAGAAGCCCCCATTCCCT
CCTTCTGTCCTCTGCTATGGGGCCTAGAGCAGCAGGAATCCAAAAGCAGTTTAAGGACA
AGGAGGGCACAAGGTCTGGATGGAGAGCATGAGTTACCCAGCTGGAACTCTGACATAGG
TTGACAGCAGCATCCCCCATTCCCAGGTGCTCATGTCTTCCCTTCTTGTGCCTTCCCTT
GGGCACTAAGTTTGGCACAGTGGCTAGGATGTAGCATTCCTCACTGGGGCCATCTGTCA
CATCAAGAAGGGTTCATTGAG

gtgtaatgtattagagcaaatcctcttgattaggcttgagaatggagccatggagcccc
atttttttcccacccttcatgcagtagtgtttaattaaatatttaaaatatttaatgcc
ctgcacaggcatcatttaattggaatgaacaactgctaactgctggcacagggctctag
aaggccccagatatcagtaatttaccactgtttgcttgctcttgggataggaaggatcc
ggggatcctagaggaggagctagggcagttgggtgctggaggaggcacatgggggctca
gcacagccacttgtttgccagctggtggagcagtgtggaactcgccttcttgggaggaa
gaaacacgtctccagacttccataacaaagtacccagagttgctgggctagttacagtt
ccaatgaccattcctccccagcaggataagcccagggccccaccctacctgggtccccc
ttctcgccccgagggccctctctcccatcccgtccatcgcgaccaggcaggccactctC
CACTGAGCTACACATGACCAGGGTGCAAGCACTGGGC[ga]TTGTTCTGTGGGAGTAGG
TCTTCATTTCTGCTTCCAGGtagcccaggggctgtgtgagcaggaccagtgcagagagg
aggaagagcagcatggcctggagaggtgaacagaaagagaaaagacatgcttatgcttc
atggacatggtttagggcttggctcagcttctagaggtgacaagaagcccccattccct
ccttctgtcctctgctatggggcctagagcagcaggaatccaaaagcagtttaaggaca
aggagggcacaaggtctggatggagagcatgagttacccagctggaactctgacatagg
ttgacagcagcatcccccattcccaggtgctcatgtcttcccttcttgtgccttccctt
gggcactaagtttggcacagtggctaggatgtagcattcctcactggggccatctgtca
catcaagaagggttcattgag

FIGURE 23tttttt

67379
ATGGCACCTGCCCTTTGGCACCCCAAGGTGGAGCCCCCAGCGACCTTCCCCTTCCAGCT
GAGCATTGCTGTGGGGGAGAGGGGGAAGACGGGAGGAAAGAAGGGAGTGGTTCCATCAC
GCCTCCTCACTCCTCTCCTCCCGTCTTCTCCTCTCCTGCCCTTGTCTCCCTGTCTCAGC
AGCTCCAGGGGTGGTGTGGGCCCCTCCAGCCTCCTAGGTGGTGCCAGGCCAGAGTCCAA
GCTCAGGGACAGCAGTCCCTCCTGTGGGGGCCCCTGAACTGGGCTCACATCCCACACAT
TTTCCAAACCACTCCCATTGTGAGCCTTTGGTCCTGGTGGTGTCCCTCTGGTTGTGGGA
CCAAGAGCTTGTGCCCATTTTTCATCTGAGGAAGGAGGCAGCAGAGGCCACGGGCTGGT
CTGGGTCCCACTCACCTCCCCTCTCACCTCTCTTCTTCCTGGGACGCCTCTGCCTGCCA
GCTCTCACTTCCCTCCCCTGACCCGCAGGGTGGCTGCGTCCTTCCAGGGCCTGGCCTGA
GGGCAGGGGTGGTTTGCTCCC[ct]CTTCAGCCTCCGGGGGCTGGGGTCAGTGCGGTGC
TAACACGGCTCTCTCTGTGCTGTGGGACTTCCAGGCAGGCCCGCAAGCCGTGTGAGCCG
TCGCAGCCGTGGCATCGTTGAGGAGTGCTGTTTCCGCAGCTGTGACCTGGCCCTCCTGG
AGACGTACTGTGCTACCCCCGCCAAGTCCGAGAGGGACGTGTCGACCCCTCCGACCGTG
CTTCCGGTGAGGGTCCTGGGCCCCTTTCCCACTCTCTAGAGACAGAGAAATAGGGCTTC
GGGCGCCCAGCGTTTCCTGTGGCCTCTGGGACCTCTTGGCCAGGGACAAGGACCCGTGA
CTTCCTTGCTTGCTGTGTGGCCCGGGAGCAGCTCAGACGCTGGCTCCTTCTGTCCCTCT
GCCCGTGGACATTAGCTCAAGTCACTGATCAGTCACAGGGGTGGCCTGTCAGGTCAGGC
GGGCGGCTCAGGCGGAAGAGC

atggcacctgccctttggcaccccaaggtggagcccccagcgaccttccccttccagct
gagcattgctgtgggggagagggggaagacgggaggaaagaagggagtggttccatcac
gcctcctcactcctctcctcccgtcttctcctctcctgcccttgtctccctgtctcagc
agctccaggggtggtgtgggcccctccagcctcctaggtggtgccaggccagagtccaa
gctcagggacagcagtccctcctgtgggggcccctgaactgggctcacatcccacacat
tttccaaaccactcccattgtgagcctttggtcctggtggtgtccctctggttgtggga
ccaagagcttgtgcccatttttcatctgaggaaggaggcagcagaggccacgggctggt
ctgggtcccactcacctcccctctcacctctcttcttcctgggacgcctctgcctgcca
gctctcacttccctcccctgacccgcagggtggctgcgtccttccagggcctggcctgA
GGGCAGGGGTGGTTTGCTCCC[ct]CTTCAGCCTCCGGGGGCTGGGGtcagtgcggtgc
taacacggctctctctgtgctgtgggacttccaggcaggcccgcaagccgtgtgagccg
tcgcagccgtggcatcgttgaggagtgctgtttccgcagctgtgacctggccctcctgg
agacgtactgtgctacccccgccaagtccgagagggacgtgtcgacccctccgaccgtg
cttccggtgagggtcctgggcccctttcccactctctagagacagagaaatagggcttc
gggcgcccagcgtttcctgtggcctctgggacctcttggccagggacaaggacccgtga
cttccttgcttgctgtgtggcccgggagcagctcagacgctggctccttctgtccctct
gcccgtggacattagctcaagtcactgatcagtcacaggggtggcctgtcaggtcaggc
gggcggctcaggcggaagagc

FIGURE 23uuuuu

67382
GAGCTGGCCCGAGCTGAAGAGTTGTTGGAGCAGCAGCTGGAGCTGTACCAGGCCCTCCT
TGAAGGGCAGGAGGGAGCCTGGGAGGCCCAAGCCCTGGTGCTCAAGATCCAGAAGCTGA
AGGAACAGATGAGGAGGCACCAAGAGAGCCTTGGAGGAGGTGCCTAAGTTTCCCCCAGT
GCCCACAGCACCCTCCGGCACTGAAAATACACGCACCACCCACCAGGAGCCTTGGGATC
ATAAACACCCCAGCGTCTTCCCAGGCCAGAGAAAGTGGAAGAGACCACAAACCGCAGGC
AATTGGCAGGCAGTGGGGGAGCCAGGGCTCTGCAGTCTTAGTCCCATTCCCCTTTGATC
TCACAGCAGGCAGGGCACCCAGGCCTTATAGGAATTCACCCTGGACCATGCCCTAAAAT
AACCTCACCCCAAATACAATAAAGGGACGAAGCACTTATAGATACCACAGACACATGTG
TTTCATTTTTAGTTTTGTTAAAAAAAAATTCTGACAAATCAGAAATGGGGGTTCAGGAG
TGGTGGTGATGCAAAAGATGGAAGCCATGGGGTGGGGGCTGTCAGGGGTGGGGGCAGTA
GTGTCTCCTT[ct]ACCCCCACCCTGGTGTCCTCTCCTGAAGGACAGACGGTCACATTC
CAAAATGGGCGAGTCTTCTACCGTGTCTGTTCAACTGAGAAGAAAACGTAGCATGGTCA
GAATAAGGCATGAAAAGGGGAAAGTGAGGCAGGAACACACGGCACACATGCAGACACTG
GTGTACTGCCTGGGTTCAGAGGACGGACGTGGGGGTGAGGGAAGGGATGTAATATGATG
AGAGAAGACAGAAACCCCACATAAAGGTCAGAAAAACATCCCAACACAGCATCAAAGAC
CAGGGGGCATGAACCAGTCAAGTGTCCATTATGCATCAGATGCCCATGACCTATGTGAT
GGGATTTAGGACAAACACACTAAGGAACAGGGAGGACCTAAAGGGTTTCATGAGATCAG
TACTCACTGTAGGAGGAGATG

gagctggcccgagctgaagagttgttggagcagcagctggagctgtaccaggccctcct
tgaagggcaggagggagcctgggaggcccaagccctggtgctcaagatccagaagctga
aggaacagatgaggaggcaccaagagagccttggaggaggtgcctaagtttcccccagt
gcccacagcaccctccggcactgaaaatacacgcaccacccaccaggagccttgggatc
ataaacaccccagcgtcttcccaggccagagaaagtggaagagaccacaaaccgcaggc
aattggcaggcagtgggggagccagggctctgcagtcttagtcccattcccctttgatc
tcacagcaggcagggcacccaggccttataggaattcaccctggaccatgccctaaaat
aacctcaccccaaatacaataaagggacgaagcacttatagataccacagacacatgtg
tttcatttttagttttgttaaaaaaaaattctgacaaatcagaaatgggggttcaggag
tggtggtgatgcaaaagatggaagccatggggtgggggctgtcagGGGTGGGGGCAGTA
GTGTCTCCTT[ct]ACCCCCACCCTGGTGTCCTCTCCTgaaggacagacggtcacattc
caaaatgggcgagtcttctaccgtgtctgttcaactgagaagaaaacgtagcatggtca
gaataaggcatgaaaaggggaaagtgaggcaggaacacacggcacacatgcagacactg
gtgtactgcctgggttcagaggacggacgtgggggtgagggaagggatgtaatatgatg
agagaagacagaaaccccacataaaggtcagaaaaacatcccaacacagcatcaaagac
cagggggcatgaaccagtcaagtgtccattatgcatcagatgcccatgacctatgtgat
gggatttaggacaaacacactaaggaacagggaggacctaaagggtttcatgagatcag
tactcactgtaggaggagatg

FIGURE 23vvvvv

67386
TTTCAATCAAAAGCTGGAAGTCCACCTTACAGAAAGACAAAAAGAAACCCCTTTTTATA
TCTTAACAAAGCAATAGCTCTCAAGCAGCAGAGCATCTCGAGGAAGAAAGCTTGCCCGG
TCGCCATCCCATCATGCCAGAGCGTGCAGTGTCCACCCTTGACTACGCTGGGGAATTGC
TGATTTTTGAAAAAGCTTAACTTAACAATTTCTGATGTCTATCTTTTAGAGTTCTGTA
TGTTCCCATTTTTTATTCTTCTGAATTTTGAATTGCAAGTAGCTGTAAAATCCAATCTT
TGAGTGCATGGGGGTGGGTGTGAGGCGGGGCTCAGCTTCAACCCCCTGTCCTGTAAAGC
AGTGGCTGGTTTTTCCTGAGCCCAGCCCTGGGAGGTCGTGGTAGGTGTGGAGGCTGCAG
AGCTCCTCCAGATGCTGCCCTCGCTGTGCCTCACACCAGAGAGGATGGAAGTGGGCTCT
GGTGTCAGACTGTGGTTGAGCTGAGACAGACAAGGCCGACACAGGGCTGGGGGCCCGTG
GTCCACCAGTGGAAGTGACTGCCGAGGAAGGG[ca]GGTGAGGAGGGCGGTGTGGGAGC
TGAGGCTTCTTTTCAGCCTGGCAGCTGGCGAGGGCCAGGGAGCAGGGGAAGAGCCTGGT
CACCATGGTCCCAGAGCCCGTCTCACTTGGCTTTTCCTTTGCAGCTGAGGAGGATGAGG
GCCAGAGAGGGACTGTGTGTATGTCCTGCCTGGGGACCCACAGCCAGGTGATAGCAGAG
GTGGTTTGAAGCCCAGGCCTCCCACGCCAACCCACTGGTCTTGCTGTTTCAGCAGGGAA
GGCCGGGAGCCCTAGGAGCTGGGGAAAGGCGACTGCCCGGGTCCTGGTGACTCCCCAC
CCCCAGATCCCCAGCTGTCATCACTGGGGCAAGGACACATTAAACTGGTCCCTGTGGGT
CAGGTCTGAGTGGGGGAGGACCTCCCCTCCCCACTGCCTCCCACAGGGGCTTGTGATGC
AGGGTTTCAGGAACAGGGCTG

tttcaatcaaaagctggaagtccaccttacagaaagacaaaaagaaacccctttttata
tcttaacaaagcaatagctctcaagcagcagagcatctcgaggaagaaagcttgcccgg
tcgccatcccatcatgccagagcgtgcagtgtccacccttgactacgctggggaattgc
tgatttttgaaaaagcttaacttaacaatttctgatgtctatcttttagagttctgta
tgttcccattttttattcttctgaattttgaattgcaagtagctgtaaaatccaatctt
tgagtgcatgggggtgggtgtgaggcggggctcagcttcaaccccctgtcctgtaaagc
agtggctggtttttcctgagcccagccctgggaggtcgtggtaggtgtggaggctgcag
agctcctccagatgctgccctcgctgtgcctcacaccagagaggatggaagtgggctct
ggtgtcagactgtggttgagctgagacagacaaggccgacacagggctgggggcccgtg
gtccaccagTGGAAGTGACTGCCGAGGAAGGG[ca]GGTGAGGAGGGCGGTGTGGGAGC
tgaggcttcttttcagcctggcagctggcgagggccagggagcaggggaagagcctggt
caccatggtcccagagcccgtctcacttggcttttcctttgcagctgaggaggatgagg
gccagagagggactgtgtgtatgtcctgcctggggacccacagccaggtgatagcagag
gtggtttgaagcccaggcctcccacgccaacccactggtcttgctgtttcagcagggaa
ggccgggagccctaggagctggggaaaggcgactgcccgggtcctggtgactccccac
cccagatccccagctgtcatcactggggcaaggacacattaaactggtccctgtgggt
caggtctgagtgggggaggacctcccctccccactgcctcccacaggggcttgtgatgc
agggtttcaggaacagggctg

FIGURE 23wwwww

67396
TTGGTTTTTGTTGTATTCAATTCTAATTATTTATTACACAGTTACCATCCTTTGATGAG
ATGTTACTCTTCATCTGTGATTGCTTATAGTTGTTCGCGAGCTTCTGTCCATTGGTAAT
TAGAAAGTTTATTTATATCAAGTTTAATCTTCCTGTTAAAAACAGTGTTCTAATAGTCA
TCCATATTAAAATATTATATGGCAGTATTAAAAACTACAAATATTACTCTTGGGAATCA
AATCATACACTGTAGCACATCATCTTTCTTGGCAATAGTACTGCTGTTGTACACTGATG
GCCTCTAACAGAGAAGAAATCATTCCATTGAAAGAAAAGTAACTATCAAGAACAAAGTT
GGAAGTGATGCCTTAAAGCTACCGGCCCATGTCTAAATGTACTTTTGATTTTTATTTTA
TTGGTTAAGTAGAAATTATTTTTAATGTAATGACAGCCCATTAATAAATGTCTCCTCTG
TTGAAGGTAGGGTTAATTCAGTATGCCAATAATCCAAGAGTTGTGTTTAACTTGAACAC
ATATAAAACCAAAGAAGAAATGATTGTAGCAACATCCCAGACATCCCAATATGGTGGGG
ACCTCACAAACACATT[ct]GGAGCAATTCAATATGCAAGGTAAGTTTTGGTGCTAATA
GGCCAATGTTTTCATAATGTAAAACATTATATTTATGTAATAAATATGAAAAAGTAAGG
AAAAGACAAAGAAAAATAATATACCTGGTACCTAATTTAAATCAGAACTAATAAAGAAA
AAAACATCAGAGCATTCTATGTCTTGAATACTTTGAGAAGGCAGCTGGGAAAGTTAAAT
CTTTGATTTTAGGATATTTATAAGATATCACATGATATTTAAATGAATTTATGTGAAGT
AAATGAAATGAGAAGACCTTAGATTAAAACAGTAGGAAATGGGGCAATCTGTCATAATT
TGTTAATATTCATCAGAGATTCAGACAAATTGAGCTCATGGATCACTTGGTGCAAATTA
ACAAAGACCACAGAATCTTAA

ttggtttttgttgtattcaattctaattatttattacacagttaccatcctttgatgag
atgttactcttcatctgtgattgcttatagttgttcgcgagcttctgtccattggtaat
tagaaagtttatttatatcaagtttaatcttcctgttaaaaacagtgttctaatagtca
tccatattaaaatattatatggcagtattaaaaactacaaatattactcttgggaatca
aatcatacactgtagcacatcatctttcttggcaatagtactgctgttgtacactgatg
gcctctaacagagaagaaatcattccattgaaagaaaagtaactatcaagaacaaagtt
ggaagtgatgccttaaagctaccggcccatgtctaaatgtacttttgatttttatttta
ttggttaagtagaaattatttttaatgtaatgacagcccattaataaatgtctcctctg
ttgaaggtagggttaattcagtatgccaataatccaagagttgtgtttaacttgaacac
atataaaaccaaagaagaaatgattgtagcaACATCCCAGACATCCCAATATGGTGGGG
ACCTCACAAACACATT[ct]GGAGCAATTCAATATGCAAGGTAAGTTTTGGTGCTAATA
GGCCAatgttttcataatgtaaaacattatatttatgtaataaatatgaaaaagtaagg
aaaagacaaagaaaaataatatacctggtacctaatttaaatcagaactaataaagaaa
aaaacatcagagcattctatgtcttgaatactttgagaaggcagctgggaaagttaaat
ctttgattttaggatatttataagatatcacatgatatttaaatgaatttatgtgaagt
aaatgaaatgagaagaccttagattaaaacagtaggaaatggggcaatctgtcataatt
tgttaatattcatcagagattcagacaaattgagctcatggatcacttggtgcaaatta
acaaagaccacagaatcttaa

FIGURE 23xxxxx

87801
TGGATCTGCAGCTCCAGAGAAGGGCCTGGGTCAGATGTCACTGAAGCCCTATGGTGGCG
GAAAGGCGAGAAATAGTGGGTTGAGATTCCAAGTGCAATCCACTGCGGCTCCTCGCTCG
CCCTCCAGGTGGCAGCACAACCCTGCGCTTCCGAAGCCCGTTTTCTGAGCCAGACACTC
TCCACGCTCTGGGTATTTCGGCTTCTCTCTCCCCACACGCCGACCCTAGGTCGCGCACT
TTCTGCCTGGCAGAATTTGGCCGAGGATCCAAACCCGGAGCAGCCTCCAGAGAGCGTGT
CGTTCACGCGGCCAGCATATGCTCAGAGACCTCAGAGGCTCAGAGACCTCAGGGCTGGT
GGTGTGGTCGGTTGTGACCACTTGTCCCTCGGACCGGCTCCAGGAACCAACCTGGGGAA
TGTGTGTAGGGGAAGGGCGGGATAGACAGTGCCCGGAGCAGGGAGGCGCTGAAAGACAG
GACCAAGCAGCCCGGCCACCAGACCCGTTGTGGGAACGGAATTTCCTGGCCCCCAGGGC
CACACTCGCGTGGGAAGCATGTCGCGGAC[ct]CTTTAAGGCGTCATCTCCCTGTCTCT
CCGCCCCCGCCTGGGACAGGCCGGGACGCCCGGGACCTGACATTTGGAGGCTCCCAACG
TGGGAGCTAAAAATAGCAGCCCCGGGTTACTTTGGGGCATTGCTCCTCTCCCAACCCGC
GCGCCGGCTCGCGAGCCGTCTCAGGCCGCTGGAGTTTCCCCGGGGCAAGTACACCTGGC
CCGTCCTCTCCTCTCAGACCCCACTGTCCAGACCCGCAGAGTTTAAGATGCTTCTGCAG
CCCGGGATCCTAGCTGGTGGGCGGAGTCCTAACACGTGGGTGGGCGGGGCCTTTTGTTC
CAGGGACTCTTTTCTCAAAACTTCCCAGTCGGAGGCTGGCGGGAACCCGAGAGGCGTGT
CTCGCCAGCCACGCGGAGGGGCGTGGCCTCATTGGCCCGCCCCACCAACTCCAGCCAAA
CTCTAAACCCCAGGCGGAGGG

tggatctgcagctccagagaagggcctgggtcagatgtcactgaagccctatggtggcg
gaaaggcgagaaatagtgggttgagattccaagtgcaatccactgcggctcctcgctcg
ccctccaggtggcagcacaaccctgcgcttccgaagcccgttttctgagccagacactc
tccacgctctgggtatttcggcttctctctccccacacgccgaccctaggtcgcgcact
ttctgcctggcagaatttggccgaggatccaaacccggagcagcctccagagagcgtgt
cgttcacgcggccagcatatgctcagagacctcagaggctcagagacctcagggctggt
ggtgtggtcggttgtgaccacttgtccctcggaccggctccaggaaccaacctggggaa
tgtgtgtaggggaagggcgggatagacagtgcccggagcagggaggcgctgaaagacag
gaccaagcagcccggccaccagacccgttgtgggaacggaatttcctggcccccaggGC
CACACTCGCGTGGGAAGCATGTCGCGGAC[ct]CTTTAAGGCGTCATCTCCCTGTCTCT
CCGCCcccgcctgggacaggccgggacgcccgggacctgacatttggaggctcccaacg
tgggagctaaaaatagcagccccgggttactttggggcattgctcctctcccaacccgc
gcgccggctcgcgagccgtctcaggccgctggagtttccccggggcaagtacacctggc
ccgtcctctcctctcagaccccactgtccagacccgcagagtttaagatgcttctgcag
cccgggatcctagctggtgggcggagtcctaacacgtgggtgggcggggccttttgttc
cagggactcttttctcaaaacttcccagtcggaggctggcgggaacccgagaggcgtgt
ctcgccagccacgcggaggggcgtggcctcattggcccgccccaccaactccagccaaa
ctctaaaccccaggcggaggg

Figure 24

	SNP number and oligonucleotide sequence	FRET CASSETTE
	29043	
Invader	CAGTTGTTTATCTTTCGCTCCATCAACCAAGTCACAATTGGT	
Probe	CGCGCCGAGGAGTTGGGAGGGAATTTCTV	1
Probe	ATGACGTGGCAGACGGTTGGGAGGGAATTTCV	2
	34573	
Invader	GGCAGGCTTCAGTTTGGCCAGGCCA	
Probe	ATGACGTGGCAGACCTCCATGGTGGTGCTV	2
Probe	CGCGCCGAGGTTCCATGGTGGTGCTV	1
	37319	
Invader	AGCATAGTCCCAGGAATGAGGTCCCCCAAT	
Probe	ATGACGTGGCAGACGTTGCTAAGTTTTACATAGGGV	2
Probe	CGCGCCGAGGATTGCTAAGTTTTACATAGGGGV	1
	37810	
Invader	CCAACACAGATGGAGATTATGGCAGACTTGTTTT	
Probe	ATGACGTGGCAGACGTTAGAAGAGCAGCCCTV	2
Probe	CGCGCCGAGGCTTAGAAGAGCAGCCCTV	1
	37879	
Invader	GCTGCACCGCCTCATCAATCCCAACTTCTC	
Probe	CGCGCCGAGGTCGGCTATCAGGACGV	1
Probe	ATGACGTGGCAGACACGGCTATCAGGACGV	2
	41064	
Invader	GCATGACAGGAAGACAGGGTGTGAGGTTGGAT	
Probe	CGCGCCGAGGAGGAGAGAGGCTGTAGV	1
Probe	ATGACGTGGCAGACGGGAGAGAGGCTGTAGV	2
	41164	
Invader	ACTGCTACTGTCTGTGCTGTGCTGGGCT	
Probe	ATGACGTGGCAGACGCAGAGCTGGACACCV	2
Probe	CGCGCCGAGGACAGAGCTGGACACCV	1
	41365	
Invader	GGTCTCTCTGGACAGCACACTGCACCAAGTAT	
Probe	CGCGCCGAGGAGCCCACCAAAAACGV	1
Probe	ATGACGTGGCAGACGGCCCACCAAAAACGV	2
	41408	
Invader	TCCTATGCCCAAGTTCTCTGATCATCCTCAAAAGAAGACAGT	
Probe	CGCGCCGAGGACTTCCATCCCAGAGGV	1
Probe	ATGACGTGGCAGACCCTTCCATCCCAGAGGV	2

Figure 24

	41409	
Invader	CTGCCRTGCCCTTCCTGGCCCAC	
Probe	CGCGCCGAGGTCCCTAAACCTAAATTCAAATCTV	1
Probe	ATGACGTGGCAGACGCCCTAAACCTAAATTCAAATCV	2
	41420	
Invader	GCTGCAGAGATGTGTCCTCCCACAGAGGAGT	
Probe	ATGACGTGGCAGACCTGAAACCACCAAGGAGV	2
Probe	CGCGCCGAGGATGAAACCACCAAGGAGV	1
	41452	
Invader	GCCTCTGGTTTCTGTCTACTCCAACGTCCACGT	
Probe	CGCGCCGAGGCGCATAGATACAGGCATCV	1
Probe	ATGACGTGGCAGACGGCATAGATACAGGCATCV	2
	41475	
Invader	GTCCGTGGGGTTTTTGCTGTGCGGAT	
Probe	ATGACGTGGCAGACGTGGAAAGTACAAGGCTCV	2
Probe	CGCGCCGAGGCTGGAAAGTACAAGGCTCV	1
	41541	
Invader	CAGAAGGCTGCAGCCTCACAATGCAGGT	
Probe	CGCGCCGAGGATGACTGGGTCCCCV	1
Probe	ATGACGTGGCAGACGTGACTGGGTCCCCV	2
	41646	
Invader	CCCAAATTTGATCCACTGTAACCGTGCGTACACAGT	
Probe	ATGACGTGGCAGACCACCGTTGCAACAACAV	2
Probe	CGCGCCGAGGAACCGTTGCAACAACAV	1
	44907	
Invader	GAGAGTTGCTCAAGGTAACACAGTGGTAAGTGACGGT	
Probe	ATGACGTGGCAGACGGCCAGGAACTAGACTV	2
Probe	CGCGCCGAGGAGCCAGGAACTAGACTCV	1
	44927	
Invader	GCAGTCAGTAGCAGCAGCTTGAGTGGCAGA	
Probe	ATGACGTGGCAGACCGGTTCTCAAACCTGGV	2
Probe	CGCGCCGAGGTGGTTCTCAAACCTGGAV	1
	45046	
Invader	CCCTCTGGAAGGATGGCTMATTTGCACACA	
Probe	ATGACGTGGCAGACCTGAGGCTTTCCTGATGV	2
Probe	CGCGCCGAGGTTGAGGCTTTCCTGATGAV	1

Figure 24

	45076	
Invader	GCGAGGCCGAGCCCCTCCTAGTGT	
Probe	ATGACGTGGCAGACGTTCCGGACCTTGCTV	2
Probe	CGCGCCGAGGCTTCCGGACCTTGCTV	1
	45080	
Invader	ACAAACCTTTTAGTTTACTCTGCAGTTAATCCCACTGAT	
Probe	ATGACGTGGCAGACGAAGTAGTGGGCTCCAV	2
Probe	CGCGCCGAGGAAAGTAGTGGGCTCCAAV	1
	47850	
Invader	TGTATGTTGGCCTCCTTTGCTGCCCTCACT	
Probe	ATGACGTGGCAGACGATCTCTTCCTGTGACACV	2
Probe	CGCGCCGAGGAATCTCTTCCTGTGACACV	1
	47852	
Invader	GCCCAGAGCGGGAGACAGCGACA	
Probe	ATGACGTGGCAGACCGACTTGGATATCAGGTACV	2
Probe	CGCGCCGAGGTGACTTGGATATCAGGTACTV	1
	47855	
Invader	TCGTGGTCCGGCGCATGGCTTCA	
Probe	CGCGCCGAGGTATTGGGTGCCAGCAV	1
Probe	ATGACGTGGCAGACCATTGGGTGCCAGCV	2
	47857	
Invader	GTGATCATTCTGATGGTGTGGATTGTGTCAGGCCTTAA	
Probe	ATGACGTGGCAGACCCTCCTTCTTGCCCAV	2
Probe	CGCGCCGAGGTCTCCTTCTTGCCCATTV	1
	47867	
Invader	AGCGACACCTTCACGTTGTCCTGGACCT	
Probe	ATGACGTGGCAGACGCCGTCTGGTTGTTCV	2
Probe	CGCGCCGAGGACCGTCTGGTTGTTCCV	1
	47868	
Invader	GCGGAGCCAAAGGACCGAGCAGGC	
Probe	CGCGCCGAGGTTTAATCCCACAGCCAGV	1
Probe	ATGACGTGGCAGACGTTAATCCCACAGCCAGV	2
	47877	
Invader	GCGTGTCCTCCAGGGTGAACATGTCCCT	
Probe	ATGACGTGGCAGACGCTGGACCTGTGTGAV	2
Probe	CGCGCCGAGGACTGGACCTGTGTGAAV	1

Figure 24

	47878	
Invader	GCATTTGATTGCAGAGCAGCTCCGAGTCCT	
Probe	CGCGCCGAGGATCCAGAGCTTCCTGCV	1
Probe	ATGACGTGGCAGACGTCCAGAGCTTCCTGCV	2
	47880	
Invader	GAACAGCTTCACCACGGCGGTCATGTT	
Probe	ATGACGTGGCAGACGCTTCTGTCCCCTGGV	2
Probe	CGCGCCGAGGACTTCTGTCCCCTGGV	1
	47885	
Invader	AACCCATAGTTAAGAACGTGGGGTGAGGTACCGC	
Probe	CGCGCCGAGGTCCTGCCCTTTGGCV	1
Probe	ATGACGTGGCAGACACCTGCCCTTTGGCV	2
	47887	
Invader	GCTGGAGTGTGCCCAATGCTATATGTCAGTTGAGT	
Probe	ATGACGTGGCAGACGTTCTAAGACTTGGAAGCCV	2
Probe	CGCGCCGAGGATTCTAAGACTTGGAAGCCV	1
	47892	
Invader	GGGAACAATCACCTTTTCTCTTTGCCTTTCATACTGCTTTAGACT	
Probe	CGCGCCGAGGACCTCACTGCTTCCTAAV	1
Probe	ATGACGTGGCAGACCCCTCACTGCTTCCTAAV	2
	47897	
Invader	TTTGTTCCGGACATCATGTGTATCCCAACCTACCAAAAT	
Probe	CGCGCCGAGGAAGTCCTTTCCAGGTAAGGV	1
Probe	ATGACGTGGCAGACGAGTCCTTTCCAGGTAAGV	2
	47899	
Invader	TTTGTGCAGTGGTTGATGAATACCAACAGGAACAGGTAAT	
Probe	CGCGCCGAGGAAGTCTAAGCCTGGCTV	1
Probe	ATGACGTGGCAGACGAGTCTAAGCCTGGCTV	2
	47901	
Invader	CAGGCTCAGGTTGTGGTGACACTGGTCACA	
Probe	CGCGCCGAGGTGTAGAGCTTCCACTTCTV	1
Probe	ATGACGTGGCAGACCGTAGAGCTTCCACTTCTV	2
	47905	
Invader	TGGCCCTGTGACTATGGCTCTGGCACA	
Probe	ATGACGTGGCAGACCACTAGGGTCCTGGCV	2
Probe	CGCGCCGAGGTACTAGGGTCCTGGCCV	1

Figure 24

	47906	
Invader	CCCATCCTGACCACCATCCGCCGAATCT	
Probe	CGCGCCGAGGAGCCTCTTCAATAGCAGTV	1
Probe	ATGACGTGGCAGACGGCCTCTTCAATAGCAGV	2
	47909	
Invader	GGAGTCAAGACCCAGATGTCCCCTGACTTGTT	
Probe	ATGACGTGGCAGACGTCACACAAGGAGTCTTCV	2
Probe	CGCGCCGAGGATCACACAAGGAGTCTTCAV	1
	47910	
Invader	CGACTGTCCAGTTAAATGCATCAGAAGTGTTAGCTTCTCCT	
Probe	ATGACGTGGCAGACGGAGTTAAAGTCATTACTGTAGAV	2
Probe	CGCGCCGAGGAGAGTTAAAGTCATTACTGTAGAGV	1
	47915	
Invader	GAGACACCTCCCACTCGTCCGGCAA	
Probe	CGCGCCGAGGTGTACACAGAGCATGGAV	1
Probe	ATGACGTGGCAGACCGTACACAGAGCATGGV	2
	47917	
Invader	CCAAGGCTGATGACATTGTTGGCCCTGTGT	
Probe	CGCGCCGAGGACGCATGAAATCTTTGAGAV	1
Probe	ATGACGTGGCAGACGCGCATGAAATCTTTGAGV	2
	47920	
Invader	CACTCCCAAATTCAATATTGACATATTCCCCCGGGCA	
Probe	CGCGCCGAGGTCTTGGGCTCTGGAGV	1
Probe	ATGACGTGGCAGACCCTTGGGCTCTGGAGV	2
	47924	
Invader	CGCCTGGCAGAGGACCCTGCCT	
Probe	CGCGCCGAGGAAGCCCAGGTACCGV	1
Probe	ATGACGTGGCAGACGAGCCCAGGTACCGV	2
	47925	
Invader	CCGTGCAGAGTGGTGTGGGCACTTTGAA	
Probe	CGCGCCGAGGTGGTGTTGCCAAACTTGV	1
Probe	ATGACGTGGCAGACCGGTGTTGCCAAACTTV	2
	47932	
Invader	GGTTCTCCCGAGAGGTAAAGAACAAAGACTTCAAAGACACTTC	
Probe	CGCGCCGAGGTCTTCACTGGTCAGCTCV	1
Probe	ATGACGTGGCAGACGCTTCACTGGTCAGCV	2

Figure 24

	47941	
Invader	TGTTGAACAGTCTTCAAGGTGGGATCGTAATAATGGCAAAAGT	
Probe	CGCGCCGAGGACCTCACCAAGAATTTGGV	1
Probe	ATGACGTGGCAGACGCCTCACCAAGAATTTGGV	2
	47950	
Invader	CAGCAATTTCCTCAAAAGACTTTCCTTTGGTTTCTGGAACTTTAAAAAATGTT	
Probe	ATGACGTGGCAGACGAACAGGGTAAAGGCCV	2
Probe	CGCGCCGAGGAAACAGGGTAAAGGCCAV	1
	47960	
Invader	GGCCCAGAAGACCCCCCTCGGAATCT	
Probe	CGCGCCGAGGAGAGCAGGGAGGATGV	1
Probe	ATGACGTGGCAGACGGAGCAGGGAGGATGV	2
	47963	
Invader	CTCCATCCGCATCGGCCTCTATGACTCCT	
Probe	CGCGCCGAGGATCAAGCAGGTGTACACV	1
Probe	ATGACGTGGCAGACGTCAAGCAGGTGTACACV	2
	47972	
Invader	GGACACTGGTCGGCAATCCTCAGCACAGT	
Probe	ATGACGTGGCAGACGACGCCACTTCCCAV	2
Probe	CGCGCCGAGGCACGCCACTTCCCAV	1
	47980	
Invader	CACCAGGCTGCCTTGGCCACAGAAA	
Probe	CGCGCCGAGGTACTTACTGAAATGCCCTTGV	1
Probe	ATGACGTGGCAGACCACTTACTGAAATGCCCTTV	2
	47981	
Invader	GCCTCTGACCCCATGGCAGGGGT	
Probe	CGCGCCGAGGACAGAGTATTTGAGCAGCV	1
Probe	ATGACGTGGCAGACGCAGAGTATTTGAGCAGCV	2
	47986	
Invader	GCTGGGGCCCCACTGCCCAT	
Probe	CGCGCCGAGGATGTCACCTTGGATGGCV	1
Probe	ATGACGTGGCAGACGTGTCACCTTGGATGGV	2
	47987	
Invader	GTTCATCTTTGGTTTTGTGGGCAACATGCTGGTCT	
Probe	CGCGCCGAGGATCCTCATCTTAATAAACTGCAV	1
Probe	ATGACGTGGCAGACGTCCTCATCTTAATAAACTGCAV	2

Figure 24

47994
Invader GCTCCACTTTCAACTTGTCCCCCTCCAGCT
Probe ATGACGTGGCAGACGTCACCTGGGAGGCV 2
Probe CGCGCCGAGGATCACCTGGGAGGCV 1

47997
Invader [illegible]
Probe CGCGCCGAGGATCAACAAGATGCTGGATAAV 1
Probe ATGACGTGGCAGACGTGAACAAGATGCTGGATAV 2

48083
Invader GTCCTGTCTCTGCAAATAATGATGCTTTCGAAGTTTCAGTTGAACA
Probe CGCGCCGAGGTGTCCCTCGCGAAAAV 1
Probe ATGACGTGGCAGACCGTCCCTCGCGAAV 2

48088
Invader GCCATCTCCTTCTTTGCGCTCCCAGCT
Probe CGCGCCGAGGAGTAGGTGCCCCGTV 1
Probe ATGACGTGGCAGACGGTAGGTGCCCCGV 2

48094
Invader GGCAGGATGAAAACACTTACGTCGGAGGATCTCTCT
Probe ATGACGTGGCAGACGTTGCTTTCTGACGTACCV 2
Probe CGCGCCGAGGATTGCTTTCTGACGTACCV 1

48095
Invader GGAGCCAAGCACTGCTCCTCCCACT
Probe ATGACGTGGCAGACGGCCAGCATGAGGCV 2
Probe CGCGCCGAGGAGCCAGCATGAGGCV 1

48097
Invader CGCGTTCAGTCCGTGCATGCGGTTCT
Probe ATGACGTGGCAGACCGCTCCCGGGCV 2
Probe CGCGCCGAGGAGCTCCCGGGCV 1

53412
Invader TGGATTATCTAAATGAAACACAGCAGCTTACTCCAGAGT
Probe CGCGCCGAGGATCAAGTCCAAGGCCAV 1
Probe ATGACGTGGCAGACGTCAAGTCCAAGGCCAV 2

53415
Invader CGGCTTGCAGACACCGTGGAAGGTTCTA
Probe ATGACGTGGCAGACCCTGGGACTGCTGGV 2
Probe CGCGCCGAGGTCTGGGACTGCTGGV 1

Figure 24

	53418	
Invader	CCATGGGGTCCCATGCTGGCAGGATAAA	
Probe	CGCGCCGAGGTGGGTTCCTGCTCTAACV	1
Probe	ATGACGTGGCAGACCGGGTTCCTGCTCTAAV	2
	53433	
Invader	CTCCCTGCAGGTCACAGTCACCACCATCT	
Probe	CGCGCCGAGGAGCTATGGGGACAAGGV	1
Probe	ATGACGTGGCAGACGGCTATGGGGACAAGGV	2
	53444	
Invader	GCCTGGTCCCCAAGGTAGGGGCT	
Probe	ATGACGTGGCAGACCAGGTCGAGGTAGCAGV	2
Probe	CGCGCCGAGGAAGGTCGAGGTAGCAGV	1
	53445	
Invader	CCCTACCTTGGGGACCAGGCCCTTGA	
Probe	ATGACGTGGCAGACCGCTGTGGAACCAGV	2
Probe	CGCGCCGAGGTGCTGTGGAACCAGGV	1
	53459	
Invader	GGGAGGACAATCCTGTGGAAAGGAAGGTTTTTATAATGTGTTT	
Probe	ATGACGTGGCAGACCTGAGAAGGAGGGTGACV	2
Probe	CGCGCCGAGGATGAGAAGGAGGGTGACV	1
	53464	
Invader	CCTGTCTGTATCCAGCTTTGCAGTTGGTGGAATGAA	
Probe	ATGACGTGGCAGACCTGCATCATTCTTTGGTGV	2
Probe	CGCGCCGAGGTTGCATCATTCTTTGGTGGV	1
	53470	
Invader	GGAAAGAAGAAAGAGCAGAGGAGGGAGATTGGAAGTAGAAATGT	
Probe	CGCGCCGAGGATGAATGCAGAGGCAAAAV	1
Probe	ATGACGTGGCAGACCTGAATGCAGAGGCAAAV	2
	53475	
Invader	GGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAT	
Probe	CGCGCCGAGGATTACCCATCTCGAAAAGAAGV	1
Probe	ATGACGTGGCAGACGTTACCCATCTCGAAAAGAAV	2
	53478	
Invader	TCCCACCCCCACTGGACTCACCACT	
Probe	ATGACGTGGCAGACGTGATGGCAGGTGAAGV	2
Probe	CGCGCCGAGGATGATGGCAGGTGAAGV	1

Figure 24

	53488	
Invader	GGTGCCGGCAGGCAAGATAGACAGCT	
Probe	ATGACGTGGCAGACGGTGGAGTAGAAGAGCTV	2
Probe	CGCGCCGAGGAGTGGAGTAGAAGAGCTGV	1
	53503	
Invader	GGTTCAGTCCACATAATGCATTTTCTCCTTCAATTCTGAAAAGTAGCTAAC	
Probe	CGCGCCGAGGTGCTCATTTGGTAGTGAAGV	1
Probe	ATGACGTGGCAGACGGCTCATTTGGTAGTGAAGV	2
	53522	
Invader	CGGCCACTGAGGGAGAAGGCCACT	
Probe	ATGACGTGGCAGACGGACGTGATGCCGV	2
Probe	CGCGCCGAGGAGACGTGATGCCGCV	1
	53528	
Invader	GGGTCTCCACCACGGCTTTCTGGTGGT	
Probe	ATGACGTGGCAGACGCCGCCTCCTCAGV	2
Probe	CGCGCCGAGGACCGCCTCCTCAGV	1
	53529	
Invader	CTGAGCCATGGTGGCCATGAAGGGGA	
Probe	CGCGCCGAGGTTCTGGGTCACATGGCV	1
Probe	ATGACGTGGCAGACCTCTGGGTCACATGGCV	2
	53530	
Invader	GGTGCCTTCTGATGGGGACGTGTCTGCT	
Probe	ATGACGTGGCAGACGCCAGGAGAGAAGGGV	2
Probe	CGCGCCGAGGACCAGGAGAGAAGGGAV	1
	53540	
Invader	CTGCCTTGTACCAGCATTACAAATAATCCAGCCACAAAT	
Probe	ATGACGTGGCAGACGTAAATGCTTTTCATTTCTGCTV	2
Probe	CGCGCCGAGGATAAATGCTTTTCATTTCTGCTV	1
	53544	
Invader	ACCAACGTTGACATGCACGTCCAGAATTGAGGT	
Probe	ATGACGTGGCAGACGGAGGTTGCCTTTGCV	2
Probe	CGCGCCGAGGAGAGGTTGCCTTTGCTV	1
	53551	
Invader	ACACTAAGGTCTCATCAGGGTTTGGGTGGCAT	
Probe	ATGACGTGGCAGACGAAGGAATGGAACCAGGV	2
Probe	CGCGCCGAGGAAAGGAATGGAACCAGGV	1

Figure 24

	53553	
Invader	CCTAGATGCCCTGCAGAATCCTTCCTGTTACGGA	
Probe	ATGACGTGGCAGACCCCCCCTCCCTGAV	2
Probe	CGCGCCGAGGTCCCCCTCCCTGAAV	1
	53557	
Invader	GCACTGGCCACCCGGGACGCT	
Probe	CGCGCCGAGGCCCCCCAAGGAAGGV	1
Probe	ATGACGTGGCAGACGCCCCCAAGGAAGGV	2
	54762	
Invader	CAGGGGTGGATGGTCTCTCACTCCCCT	
Probe	ATGACGTGGCAGACGGGCCTGTATTCAGTCV	2
Probe	CGCGCCGAGGCGGCCTGTATTCAGTCV	1
	54777	
Invader	TGGTGACCCTGCCCAGATGTGAAGTGTACAT	
Probe	CGCGCCGAGGACTCTGTGTTGGGGAGV	1
Probe	ATGACGTGGCAGACGCTCTGTGTTGGGGAV	2
	54874	
Invader	CTCAGCCTTAAAAAGACCTCCAGGGCTTGATGCA	
Probe	CGCGCCGAGGTGGTATGTTGTCAGGCTV	1
Probe	ATGACGTGGCAGACCGGTATGTTGTCAGGCV	2
	54915	
Invader	GCTGGAGGAGGCTATGAGAAGTGAGGTTTGCAT	
Probe	CGCGCCGAGGAGAAGAAAGAGGGGCAGV	1
Probe	ATGACGTGGCAGACGGAAGAAAGAGGGGCAV	2
	55560	
Invader	CAATGGGACGCCATAGAGGGCTTTTGAGTAGACATATT	
Probe	CGCGCCGAGGATCAGTGTAGAAGGGTGAAV	1
Probe	ATGACGTGGCAGACGTCAGTGTAGAAGGGTGAV	2
	66865	
Invader	ACACATGTGTTTCATTTTTAGTTTTGTTAAAAAAAAATTCTGACAAATCAT	
Probe	ATGACGTGGCAGACGAAATGGGGGTTCAGGV	2
Probe	CGCGCCGAGGAAAATGGGGGTTCAGGAV	1
	66902	
Invader	GGAGGAGAGCAGGCATTGGGCTAAGGAGC	
Probe	ATGACGTGGCAGACGGGGCAGTGGGCV	2
Probe	CGCGCCGAGGTGGGCAGTGGGCV	1

Figure 24

	66934	
Invader	GGGACCCATTCCTGTGTAATACAATGTCTGCACCAT	
Probe	CGCGCCGAGGATGCTAATAAAGTCCTATTCTCTTV	1
Probe	ATGACGTGGCAGACGTGCTAATAAAGTCCTATTCTCTV	2
	66938	
Invader	GACAGAGGCTTCTAGAGGGGCCAGCAGT	
Probe	ATGACGTGGCAGACGTTTGGGGAGACTTGGV	2
Probe	CGCGCCGAGGATTTGGGGAGACTTGGGV	1
	66984	
Invader	CCTCCAGGCTGGCCCCCTAGATTGCT	
Probe	ATGACGTGGCAGACGTCTGCTCCTGGCAV	2
Probe	CGCGCCGAGGATCTGCTCCTGGCATV	1
	66985	
Invader	TGGACTCTGAGCCCCACCTGCGAGA	
Probe	ATGACGTGGCAGACCCCCTAGAATCACAGAGAV	2
Probe	CGCGCCGAGGTCCCTAGAATCACAGAGAGV	1
	67309	
Invader	GGGTGCTGTCCACACTGGCTCCCT	
Probe	ATGACGTGGCAGACGTCAGGGAGCAGCCV	2
Probe	CGCGCCGAGGATCAGGGAGCAGCCV	1
	67320	
Invader	TCATGAACAGCAAAGGCGTGAGCCTCTTCGT	
Probe	CGCGCCGAGGACATCATCAACCCTGAGAV	1
Probe	ATGACGTGGCAGACGCATCATCAACCCTGAGV	2
	67325	
Invader	GGTGGGGCTGGGCTGCTAGGGT	
Probe	ATGACGTGGCAGACGATCCAGATGGCATGTGV	2
Probe	CGCGCCGAGGAATCCAGATGGCATGTGV	1
	67329	
Invader	CTTGGGCCACGGAGGGCAATGACCT	
Probe	CGCGCCGAGGAAGGGTGCCCCTGV	1
Probe	ATGACGTGGCAGACGAGGGTGCCCCTGV	2
	67340	
Invader	AGTGTGGTGCAGAAAACCCTTCACCCCCT	
Probe	ATGACGTGGCAGACGTGTCAAAAGGAGCTGACV	2
Probe	CGCGCCGAGGATGTCAAAAGGAGCTGACV	1

Figure 24

	67342	
Invader	GGTCTCTACCTTGGGTGCTGTTCTCTGCCTCT	
Probe	CGCGCCGAGGAGGAGCTCTCTGTCAATTV	1
Probe	ATGACGTGGCAGACGGGAGCTCTCTGTCAAV	2
	67356	
Invader	GTAGGGAGAAGTGCGGCACAGCTAAAGGAGT	
Probe	ATGACGTGGCAGACGAGCGCCTGCACCV	2
Probe	CGCGCCGAGGAAGCGCCTGCACCV	1
	67363	
Invader	GCTACGTTTTCTTCTCAGTTGAACAGACACGGTAGAAGACTCC	
Probe	ATGACGTGGCAGACGCCCATTTTGGAATGTGAV	2
Probe	CGCGCCGAGGTCCCATTTTGGAATGTGACV	1
	67371	
Invader	CATGACCAGGGTGCAAGCACTGGGCT	
Probe	ATGACGTGGCAGACGTTGTTCTGTGGGAGTAGV	2
Probe	CGCGCCGAGGATTGTTCTGTGGGAGTAGGV	1
	67382	
Invader	GGAGAGGACACCAGGGTGGGGGTT	
Probe	ATGACGTGGCAGACGAAGGAGACACTACTGCCV	2
Probe	CGCGCCGAGGAAAGGAGACACTACTGCCV	1
	87801	
Invader	GCGGAGAGACAGGGAGATGACGCCTTAAAGT	
Probe	ATGACGTGGCAGACGGTCCGCGACATGV	2
Probe	CGCGCCGAGGAGTCCGCGACATGCV	1

Figure 25a

SNP number	forward primer sequence	reverse primer sequence
29043	cgggctacccatgggaca	gtcttctggtattaagccgtaatttgca
34573	cagtntcaccagctgtggtagaacca	aagaggagcatcactgtgaccca
37319	tcccttcctcagattatattcatcccagaaa	tcaaccccctgacattatcttggatcc
37474	cactccccaacatctcatttatttttcaca	gtcatggcaatcagttggtgaaagca
37810	tcttctttagactgccacgaggaaaaa	gggagatgaggtactcactagttaaca
37879	ccctgaggaactcacgcagac	gcacctctttgcgcaggaaga
41064	agtggtggcgctctcacaaa	catttgttcaggcattacagtaaaatgcca
41164	cagggacaatcccatcccca	gtgaattgtccatgatgagagccactac
41365	tgtcccagactgggtcagca	gaatgaagaaggtactgtgggcca
41408	ctggaaacttctgccagattgttccta	caaaggactccttgtcccctagaa
41409	ggtcctttgcgcaaaggca	tttcagctcccctcctccca
41420	gtctgccttctcacagctttcc	aggtgtaacttgagtctctgcctaac
41452	agctgctgggccagca	caagctttaaaggcagtcgacattaaga
41475	gccagggatctagggctcc	cccgtcctacccagacga
41541	tcctgctgacattccgcca	ggtgcaccacccattccca
41646	gcaatcctggttaaggacttaagaattgtca	acaaaccaacgccacttcctaac
44907	cagcgtggcagagtggc	actctacgatgtgggcatttcagaga
44927	gcgcacctgtccgtagca	gccccaacaagctctcactca
45046	gaagagggtgaatactataaaaatagacttaccttcc	ttggctcaaatcgtgggataattctaagaaa
45076	ccaccgccacctccga	gacccagcagaggtccgaa
45080	tttcaaaactatcaggacctttatcattcataggaaataa	ttttaagatacctttccaagttctccctca
47850	gacttacattaggcagtgactcgatgaa	cattgctgagaacattgcctatggaga
47852	ccctggagggagttgaccc	gctcagtatgcctttcctcccc
47855	gcaacccgggaacggca	tcgtccctttcctgcgtgac
47857	cctgctgaccaagaataaggccc	cattggcatagcagttgatggcttcc
47867	catcagcatcggttctgccc	ggcgatgctcagcccgaa
47868	tccctctgtttctttccctcacaga	ggttgctgaagttgtgtgtgatcac

Figure 25b

47877	ttccttagattcttctttggagcagaataaaaga	cacaccatgtgaggtcatcagcaa
47878	gctccagggaggactcacca	catgacctcagggatgcccaca
47880	ggcccgaacatagtaattcctggtaaa	cgagtgggagaggccca
47885	tgtattacataaaccctactccaaacaaatgca	gccagcaaacacatccaggaaca
47887	cgtttcttccatccttccaggatttgaa	acctctctgtgctttctgtatcctca
47892	caggcaactggaactgaaaccc	ctcagcttccaagggccattca
47895	ggctggacatccacttcatccac	cgtagaaagagccgggcca
47897	agaatcggctgtctttgatgctgtaa	cttatacttttagaaaaaagaagacattatcaagatattcatttttgtca
47899	acgagcataagaacttaataatgtcaagagaaattttaga	tgactacagcaagtatctggactcca
47901	acaggtcccctccgctca	ggccaggcacaggctgaaa
47905	ccagagccactacctttgtcca	ggtgtcttaggagagaaaaaaaggtagaaaaa
47906	tgaccaaatgccctcaccttca	cacaatatgctggatgactcctcagac
47909	tggttgttgaggtccctgaatcc	cagggtccagctggagca
47910	gagaggcacccttcacaggaaa	aagaaaatacttctttgagctcaactacgaac
47915	gaaggagccctgcccca	agacccccaagggatcctcc
47917	ttcggctcctgccacatcaa	tgcctctcacttcctctccttaca
47920	gtagccagactgatcactcccaaa	gcagcagcagcagcagca
47922	gacgttgccgaagcccac	agataggcaaaccctacaacagca
47924	cacaaagcgggccctcc	ccccgaggaatacgtgctgac
47925	cagtaggctgtggtcctcatca	gcccattgtagctgaggaggac
47927	gggtgctggtctcataggtctca	agctcctacatcaccagtgagatcc
47932	ccggtttggttctcccgaga	cccaaggaggagctgctgaaga
47941	gctcaggaacttcaggattgctacc	agaaacaaagtagatgcatttgattcaagtttcttaaaa
47960	cagggtcctacacacaaatcagtca	ccttctgtctcggtttcttctcca
47963	gttcagggacctggtcactcac	ggcctgcagcgccaga
47972	agggcgttggcgttttcc	agggcttgatggcctctcaga
47980	agcctacccatcttccattcctca	gccaggccccccttaggac
47981	gcttctgcactgaaagggctca	ccacatggcctaccctccc
47986	gcctgtgcccagcagca	gcctaccctggcagccc

Figure 25c

47987	caactcctgcctccgctctac	gccaggttgagcaggtaaatgtca
47994	ccactctcctcctacactgtccc	aggcccccatcgatctccc
47997	gcaaagatggctctcttcatcatagtgaa	cgttctcacatgcatgccccc
48083	accaaaatcgaggtggccca	catcagaaagaaaaatgaatctgcaacttcaatagtca
48088	caggaccccagctgtccaa	cgggaagaccatcgcctcc
48094	gcacccctatgaagacccagaa	caaaggtcacttcaggttgaggca
48095	tccagtgttgtagccaaactgca	gtgtggtttgtttctccgcagaa
48097	ggcaccaccttgcgca	cgcctggagcgttttaaattgaga
53412	gttgaaataacattcaagttttcccttactcaagtaa	cagaatatggtcctctttgctcctaaca
53415	cagcagaaccacgggcac	ccacacgcttccctctaattggac
53418	agctggagggcagtatcactca	ggtggacaggaagcatgtccc
53433	gaggcgatggtcttcccga	ccgtggctgaccactgtcc
53444	cgagctgcggccattctca	acctggttccacagcgcaa
53445	gaccgtctgctacctcgacc	caagattcccatttggaggaacggaa
53459	agcagctaataataaaccagtaatttgggatagac	gtgactccgagggcagaca
53464	gtttatgcttatttatgaaatttgcctaccttccaa	ggcagctgctcaactaatcacca
53470	ctgtctgctcctctctcatcatcc	ggacagaagcaagtctgcagatca
53475	tctacaagaaaacatcagaaactcttcattcaataga	gaagccaagtattgacagctattcgaaga
53478	gggaagggtcaggaaagcca	cgagagcggattgagttcctcaa
53488	gagccacgagctcccaca	ggtagccctttaaaaggcctcc
53503	tgacatgttcgaaacctgtccataaagtaa	ggaaagaaaagcttttgttcagagctttagaaaa
53522	gctgatctgcttctcccacga	agtcgtcgtagccagcgaa
53528	gcagggctccttactgcagaa	cacgccacccatcctcaaaga
53529	gcacagggcgctcacc	cctaccagcagccgctca
53530	aggctcccttagatgcctgaca	cagggcgctgacaccca
53540	atttctcctctgtgtcttgaagggaac	ctgccccctcaccctac
53544	cctttcatttttcccggcacaga	gggaacttctttcccctcgca
53550	ggagtttctgtcctgggaggaaaaa	aacactcgtgaagctggcca
53551	ggccacagagcctggaga	cggcttgcctgtgcagtca

Figure 25d

53553	gccagccccccttcctttcc	acactgccaggagacacagaac
53557	ggagcagatcctggcaaagatcc	cgtactgcacaaacttgctgca
54762	ggtgtaggtagagataagaagagtgatactca	gctggtgacttgccccaga
54777	tgtcataatgcagtgggattgcca	caagctggcaatggtggacaca
54915	ttttaactctctgctgttccctcacc	actgacagggaatctccagaagtca
55560	gacctgatttgattgagagccttgaac	acaagccctggactagatgatttctaaga
66865	ggtgatgcaaaagatggaagcca	ccattttggaatgtgaccgtctgtcc
66902	gagagggtcatgcagtggca	tccggtgctccatggatgaca
66934	gccatactgcagcactttaaaggac	ctgctgtgatttatctgctgaaagctca
66938	catctaactgctccccagtcaca	gccctcggtcctccaggaa
66984	ccacccacccaggacacac	ccccaacggccaggcaaa
66985	ggacaaatgttctgggtctctaatattccaa	gggtgggacggagtccc
67309	gtttgccttaccttggaagtggac	tgctgagaagattgacaggttcatgca
67329	ggcccagccactgacca	cttcttggctgttgtttctgttccc
67330	cccgactgtgccgcca	gcccttttccaggtctgacaac
67340	cctctcaatgggtcacttggcaa	gtccaaatttctgttgggttcagtgaaa
67342	gaagggccaatagccctccc	gccccagccaagaaaggtcaa
67356	ttctcctggcctgtagggagaa	ccctcgtcacttcctctgtcc
67363	gtgtctccttcacccccacc	actttccccttttcatgccttattctgac
67371	cgaccaggcaggccac	gtcctgctcacacagcccc
67382	ggtgatgcaaaagatggaagcca	ccattttggaatgtgaccgtctgtcc
87801	cggaatttcctggccccca	gtcccggcctgtccca

Figure 26

Multiplex PCR 092701_Rxn1_Std Conditions (100mMKCl, 3mMMgCl, 10mMTris8.0, 2.5x Primers) Program (94,30s/55,44s)x50 w/hotstart

Snp number		10min		20 min		40 min		80 min		160 min			10 min		20 min		40 min		80 min		160 min		Calls in agreement	Mis-calls or No-calls	
107-plex (101)	Primer Count	Dye Call	Final Call	Dye Call	Final Call	Dye Call	Final Call	Dye Call	Final Call	Dye Call	Final Call	Genomic Call		46		58		78		77		73	94	SNP No.	7
47887	1	038	38	HET	HET	HET	HET	HET	HET	HET	HET	HET		0	HET	1	HET	1	HET	1	HET	1	1	0	0
53488	2	39	39	39	39	41	41	FAM	C	FAM	C	FAM		0		0		0	FAM	1	FAM	1	1	0	0
67329	3	FAM	C	FAM	C	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
67342	4	HET	HET	FAM	C	FAM	C	FAM	G	FAM	C	FAM	HET	0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
67371	5	53	53	HET	HET	HET	HET	HET	HET	HET	HET	HET		0	HET	1	HET	1	HET	1	HET	1	1	0	0
34573	6	FAM	C	FAM	C	FAM	C	FAM	G	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
37810	7	28	28	RED	G	RED	G	RED	G	RED	G	RED		0	RED	1	RED	1	RED	1	RED	1	1	0	0
41420	8	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
41475	9	39	39	39	39	38	38	HET	HET	HET	HET	HET		0		0		0	HET	1	HET	1	1	0	0
41541	10	FAM	G	FAM	G	FAM	G	FAM	G	FAM	G	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
44907	11	39	39	FAM	C	FAM	C	FAM	C	FAM	C	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
44927	12	RED	T	RED	T	RED	T	RED	T	RED	T	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
45046	13	HET	HET	HET	HET	HET	HET	HET	HET	53	53	HET	HET	1	HET	1	HET	1	HET	1		0	1	0	0
45076	14	39	39	39	39	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
45080	15	68	68	RED	T	53	53	RED	T	RED	T	RED		0	RED	1		0	RED	1	RED	1	1	0	0
47850	16	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
47855	17	39	39	39	39	FAM	G	FAM	G	FAM	G	FAM		0		0	FAM	1	FAM	1	FAM	1	1	0	0
47857	18	FAM	C	FAM	C	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47868	19	FAM	C	HET	HET	HET	HET	HET	HET	HET	HET	FAM	FAM	1	HET	0	HET	0	HET	0	HET	0	1	0	0
47877	20	FAM	C	FAM	C	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47878	21	RED	A	RED	A	RED	A	RED	A	HET	HET	RED	RED	1	RED	1	RED	1	RED	1	HET	0	1	0	0
47880	22	FAM	C	FAM	C	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47899	23	17	17	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47905	24	FAM	C	FAM	G	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47906	25	FAM	G	FAM	G	FAM	G	FAM	G	FAM	G	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47909	26	RED	T	RED	T	RED	T	RED	T	RED	T	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
47917	27	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
47920	28	29	29	29	29	RED	A	RED	A	RED	A	RED		0		0	RED	1	RED	1	RED	1	1	0	0
47924	29	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
47941	30	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
47960	31	39	39	37	37	FAM	G	FAM	G	FAM	G	FAM		0		0	FAM	1	FAM	1	FAM	1	1	0	0
47963	32	39	39	HET	HET	FAM	G	FAM	G	FAM	G	FAM		0	HET	0	FAM	1	FAM	1	FAM	1	1	0	0
47972	33	RED	G	RED	G	RED	G	RED	G	RED	G	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
47981	34	39	39	39	39	37	37	HET	HET	HET	HET	FAM		0		0		0	HET	0	HET	0	0	47981	1
53478	35	FAM	C	FAM	C	FAM	G	43	43	HET	HET	FAM	FAM	1	FAM	1	FAM	1		0	HET	0	1	0	0
53528	36	59	59	38	38	FAM	C	43	43	FAM	C	FAM		0		0	FAM	1		0	FAM	1	1	0	0
53529	37	FAM	G	43	43	HET	HET	HET	HET	HET	HET	HET	FAM	0		0	HET	1	HET	1	HET	1	1	0	0
53557	38	39	39	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
53530	39	39	39	41	41	FAM	C	FAM	C	FAM	C	FAM		0		0	FAM	1	FAM	1	FAM	1	1	0	0
47986	40	29	29	39	39	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
47987	41	43	43	43	43	FAM	G	43	43	HET	HET	FAM		0		0	FAM	1		0	HET	0	1	0	0
47994	42	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
47997	43	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
48088	44	19	19	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0

Figure 26

48097	45	39	39	37	37	FAM	G	43	43	43	43	FAM		0		0	FAM	1		0		0	1	0	0
53415	46	HET	HET	43	43	FAM	G	FAM	G	HET	HET	FAM	HET	0		0	FAM	1	FAM	1	HET	0	1	0	0
53418	47	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
53433	48	19	19	39	39	FAM	G	HET	HET	HET	HET	FAM		0		0	FAM	1	HET	0	HET	0	1	0	0
53444	49	FAM	G	FAM	G	FAM	G	FAM	G	43	43	FAM	FAM	1	FAM	1	FAM	1	FAM	1		0	1	0	0
53459	50	FAM	G	FAM	G	FAM	G	FAM	G	FAM	G	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
53522	51	52	52	53	53	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
53544	52	19	19	31	31	43	43	HET	HET	HET	HET	HET		0		0		0	HET	1	HET	1	1	0	0
54762	53	43	43	FAM	C	FAM	C	FAM	C	FAM	C	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
54777	54	RED	A	RED	A	RED	A	RED	A	53	53	RED	RED	1	RED	1	RED	1	RED	1		0	1	0	0
54915	55	RED	A	RED	A	RED	A	RED	A	53	53	RED	RED	1	RED	1	RED	1	RED	1		0	1	0	0
66865	56	39	39	39	39	39	39	59	59	58	68	FAM		0		0		0		0		0	0	66865	1
66934	57	29	29	28	28	RED	A	RED	A	RED	A	RED		0		0	RED	1	RED	1	RED	1	1	0	0
66984	58	FAM	G	FAM	G	FAM	G	FAM	G	FAM	G	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
66985	59	29	29	29	29	39	39	59	69	39	39	FAM		0		0		0		0		0	0	66985	1
67309	60	53	53	53	53	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
67340	61	53	53	53	53	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
67356	62	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
67363	63	FAM	G	43	43	HET	HET	HET	HET	HET	HET	FAM	FAM	1		0	HET	0	HET	0	HET	0	1	0	0
67382	64	17	17	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
87801	65	39	39	39	39	49	49	39	39	37	37	FAM		0		0		0		0		0	0	87801	1
53551	66	17	17	FAM	C	FAM	C	FAM	C	FAM	C	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
41084	67	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
41408	68	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
41452	69	39	39	29	29	58	58	RED	C	RED	C	FAM		0		0		0	RED	0	RED	0	0	41452	1
47901	70	39	39	39	39	37	37	39	39	HET	HET	FAM		0		0		0		0	HET	0	0	47901	1
47932	71	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
48094	72	FAM	G	FAM	C	FAM	C	FAM	C	FAM	C	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
48095	73	39	39	FAM	G	FAM	G	FAM	C	FAM	C	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
53412	74	29	29	28	28	28	28	RED	A	RED	A	RED		0		0		0	RED	1	RED	1	1	0	0
53475	75	39	39	39	39	39	39	47	47	FAM	G	FAM		0		0		0		0	FAM	1	1	0	0
53503	76	38	38	38	38	38	38	HET	HET	HET	HET	HET		0		0		0	HET	1	HET	1	1	0	0
53540	77	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
55560	78	RED	A	RED	A	RED	A	RED	A	RED	A	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
66938	79	39	39	HET	HET	FAM	C	FAM	C	FAM	C	FAM		0	HET	0	FAM	1	FAM	1	FAM	1	1	0	0
47867	80	53	53	53	53	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
48083	81	39	39	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
53553	82	FAM	C	FAM	C	FAM	C	FAM	C	HET	HET	FAM	FAM	1	FAM	1	FAM	1	FAM	1	HET	0	1	0	0
37879	83	FAM	T	FAM	T	FAM	T	FAM	T	FAM	T	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
41365	84	19	19	39	39	HET	HET	HET	HET	HET	HET	HET		0		0	HET	1	HET	1	HET	1	1	0	0
41646	85	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	HET	1	HET	1	HET	1	HET	1	HET	1	1	0	0
47852	86	FAM	C	FAM	C	FAM	C	HET	HET	53	53	17	FAM	0	FAM	0	FAM	0	HET	0		0	1	0	0
47885	87	39	39	39	39	FAM	T	FAM	T	FAM	T	FAM		0		0	FAM	1	FAM	1	FAM	1	1	0	0
47892	88	17	17	FAM	G	FAM	G	FAM	G	FAM	G	FAM		0	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0
47915	89	FAM	G	FAM	G	FAM	G	FAM	G	HET	HET	17	FAM	0	FAM	0	FAM	0	FAM	0	HET	0	1	0	0
47925	90	FAM	G	43	43	HET	HET	HET	HET	53	53	FAM	FAM	1		0	HET	0	HET	0		0	1	0	0
47980	91	FAM	G	FAM	G	43	43	HET	HET	HET	HET	FAM	FAM	1	FAM	1		0	HET	0	HET	0	1	0	0
37319	92	FAM	G	FAM	C	FAM	C	HET	HET	HET	HET	FAM	FAM	1	FAM	1	FAM	1	HET	0	HET	0	1	0	0
29043	93	FAM	G	FAM	G	FAM	G	FAM	G	FAM	G	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0

Figure 26

53445	94	FAM	C	FAM	C	HET	HET	HET	HET	53	53	17	FAM	0	FAM	0	HET	0	HET	0		0	1	0	0
41409	95	39	39	39	39	HET	HET	HET	HET	HET	HET	FAM		0		0	HET	0	HET	0	HET	0	0	41409	1
47897	96	39	39	38	38	HET	HET	43	43	FAM	G	FAM		0		0	HET	0	·	0	FAM	1	1	0	0
41164	97	RED	T	RED	T	RED	T	RED	T	RED	T	RED	RED	1	RED	1	RED	1	RED	1	RED	1	1	0	0
66902	98	39	39	37	37	FAM	G	43	43	FAM	G	FAM		0		0	FAM	1		0	FAM	1	1	0	0
53464	99	FAM	C	FAM	C	FAM	C	FAM	C	HET	HET	FAM	FAM	1	FAM	1	FAM	1	FAM	1	HET	0	1	0	0
53470	100	FAM	C	FAM	C	FAM	C	FAM	C	FAM	C	17	FAM	0	FAM	0	FAM	0	FAM	0	FAM	0	1	0	0
47910	101	FAM	C	FAM	C	FAM	C	FAM	C	FAM	D	FAM	FAM	1	FAM	1	FAM	1	FAM	1	FAM	1	1	0	0

Figure 27

TWT SNP#	Fam FOZ	Red FOZ	FOZ RATIO	Dye Call	Final Call	Genomic Call		TWT SNP#	Fam FOZ	Red FOZ	FOZ RATIO	Dye Call	Final Call	Genomic Call
47887	4.3	5.6	0.7	HET	HET	HET		53522	2.7	4.9	0.4	HET	HET	HET
53488	16.9	2.4	11.6	FAM	C	FAM		53544	9.7	4.9	2.2	HET	HET	HET
67329	6.5	1.5	11.0	FAM	C	FAM		54762	5.2	1.4	11.7	FAM	C	FAM
67342	4.5	1.2	19.1	FAM	G	FAM		54777	1.1	11.9	0.0	RED	A	RED
67371	16.2	32.3	0.5	HET	HET	HET		54916	1.2	10.4	0.0	RED	A	RED
34573	4.4	1.0	337.1	FAM	C	FAM		66865	1.0	1.1	0.3	59	59	FAM
37610	1.3	8.3	0.0	RED	G	RED		66934	1.1	8.0	0.0	RED	A	RED
41420	13.6	19.6	0.7	HET	HET	HET		66984	7.8	2.0	6.5	FAM	G	FAM
41475	1.3	1.3	1.0	HET	HET	HET		66985	1.1	1.2	0.5	39	39	FAM
41541	13.0	1.4	27.9	FAM	G	FAM		67309	2.9	5.4	0.4	HET	HET	HET
44907	4.2	1.2	16.2	FAM	C	FAM		67340	4.1	5.8	0.6	HET	HET	HET
44927	1.8	19.0	0.0	RED	T	RED		67356	6.0	6.6	0.9	HET	HET	HET
45046	3.7	4.7	0.7	HET	HET	HET		67363	1.8	1.2	4.7	FAM	G	FAM
45076	1.9	1.7	1.2	HET	HET	HET		67382	16.4	1.2	73.8	FAM	G	FAM
45080	1.4	11.1	0.0	RED	T	RED		67801	12.3	1.4	30.6	FAM	C	FAM
47850	3.9	3.6	1.1	HET	HET	HET		53551	6.9	1.1	45.8	FAM	C	FAM
47855	1.9	1.4	2.6	43	43	FAM		41064	1.5	5.4	0.1	RED	A	RED
47857	9.4	1.1	58.8	FAM	C	FAM		41408	1.2	20.5	0.0	RED	A	RED
47868	29.5	3.6	10.1	FAM	C	FAM		41452	1.3	1.6	0.5	HET	HET	FAM
47877	14.2	1.3	41.2	FAM	C	FAM		47901	1.6	1.4	2.1	HET	HET	FAM
47878	1.4	5.8	0.1	RED	A	RED		47932	1.3	13.8	0.0	RED	A	RED
47880	5.7	1.4	11.4	FAM	C	FAM		48094	5.9	1.3	15.5	FAM	C	FAM
47899	11.5	1.5	21.0	FAM	G	FAM		48095	12.1	1.2	62.3	FAM	C	FAM
47905	6.2	1.2	27.1	FAM	C	FAM		53412	1.5	11.2	0.1	RED	A	RED
47906	20.1	1.4	52.4	FAM	G	FAM		53475	10.4	1.3	28.2	FAM	G	FAM
47909	1.7	22.2	0.0	RED	T	RED		53503	3.9	3.9	1.0	HET	HET	HET
47917	2.4	26.7	0.1	RED	A	RED		53540	17.6	18.7	0.9	HET	HET	HET
47920	1.3	2.2	0.2	RED	A	RED		55560	1.4	10.9	0.0	RED	A	RED
47924	1.3	5.9	0.1	RED	A	RED		66938	9.3	1.6	13.5	FAM	C	FAM
47941	8.2	9.0	0.9	HET	HET	HET		47867	4.9	7.5	0.8	HET	HET	HET
47960	5.7	2.1	4.5	FAM	G	FAM		48083	4.3	1.1	23.8	FAM	G	FAM
47963	6.9	1.7	11.9	FAM	G	FAM		53553	14.3	1.4	30.5	FAM	C	FAM
47972	1.6	40.9	0.0	RED	G	RED		37879	7.5	1.8	13.1	FAM	[illegible]	FAM
47981	2.6	1.3	6.3	FAM	G	FAM		41365	2.2	2.1	1.1	HET	HET	HET
53478	27.0	1.6	39.9	FAM	C	FAM		41646	8.2	9.2	0.9	HET	HET	HET
53528	1.9	1.4	2.1	HET	HET	FAM		47852	10.2	1.4	23.6	FAM	C	17
53529	17.1	10.2	1.8	HET	HET	HET		47885	2.3	1.3	5.1	FAM	T	FAM
53557	7.6	1.3	20.9	FAM	G	FAM		47892	9.0	1.2	38.2	FAM	G	FAM
53530	5.4	1.7	6.3	FAM	C	FAM		47915	6.7	1.2	43.1	FAM	G	17
47986	1.4	1.7	0.6	HET	HET	HET		47925	16.9	3.6	6.0	FAM	G	FAM
47987	12.0	1.7	15.0	FAM	G	FAM		47930	17.0	1.6	25.1	FAM	G	FAM
47994	23.6	31.7	0.7	HET	HET	HET		37319	19.5	1.6	26.6	FAM	C	FAM
47997	1.5	8.3	0.1	RED	A	RED		29043	10.6	1.4	23.3	FAM	G	FAM
48088	6.4	1.2	27.0	FAM	G	FAM		53445	12.6	1.8	14.5	FAM	C	17
48097	3.7	1.6	4.2	FAM	G	FAM		41409	12.2	1.8	13.9	FAM	G	FAM
53415	3.6	1.6	4.0	FAM	G	FAM		47897	5.0	1.7	6.1	FAM	G	FAM
53418	1.3	38.8	0.0	RED	A	RED		41164	1.5	7.2	0.1	RED	T	RED
53433	1.3	1.2	1.3	39	39	FAM		66902	13.9	1.7	18.2	FAM	G	FAM
53444	16.2	1.7	22.7	FAM	G	FAM		53464	21.8	2.5	13.6	FAM	C	FAM
53459	5.7	1.4	11.2	FAM	G	FAM		53470	14.5	1.4	31.8	FAM	C	17
								47910	18.2	1.3	55.0	FAM	C	FAM

Invader Analysis of Highly Multiplexed PCR. Multiplex PCR was carried out under standard conditions using only 10ng of hgDNA as template. After 10 min at 95°C, Taq (2.5 units) was added to the 50ul reaction and additional 3ul of PCR carried out for 50 cycles. The PCR reaction was diluted and loaded directly onto an Invader™ MAP plate (3ul/well). An 15mM $MgCl_2$ was added to each reaction on the Invader™ MAP plate and covered with 6ul of mineral oil. The entire plate was heated to 95°C for 5 min. and incubated at 63°C for 40 min. FAM and RED fluorescence was measured on a Cytofluor 4000 fluorescent plate reader and "Fold Over Zero" (FOZ) values were calculated for each amplicon. Results from each SNP are color coded in the table above as "pass" (dark gray), "mis-call" (light pink), or "no-call" (white).

Figure 28

MAP092701_PCR1_ Optimization

Snp number	10 min		20 min		40 min.		80 min		160 min.		RED_FOZ + FAM_FOZ -2 values at a given time "m"					FOZm -2 (FOZm must be >2 at 10min or >2 and <12 at all other time points)					Calculation of Amplification Factor (F) using equation 1b. (w/ FOZ240=3.42 and a dilution factor of 0.125)					Selected Value of F	X[n]	R[n] (setting X[max] to 1.6)	Concentration of Primers in Optimized Mix (uM)	
107-plex (102)	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	10	20	40	80	160	10	20	40	80	160	10	20	40	80	160					
47837	1.25	1.375	1.75	2.5	1.7273	2.1333	8.1667	14.545	16.909	22.78	0.6	2.3	1.9	21.7	37.7	0	2.25	0	0	0	0	14260.66	0	0	0	14260.6	0.8775	1.82	0.046	
53488	1.1176	1.1	1.1429	1.0833	2.02	1.2500	2.3913	1.2381	3.6317	1.3443	0.2	0.2	1.3	1.6	3.0	0	0	0	0	3.035992	0	0	0	0	300.6594	300.0	-0.1208	12.00	0.300	Yellow -Low End Correction
67329	4.0667	1.125	10.056	1.2	16.182	1.1667	10.31	1.6667	9.5	2.561	3.8	9.3	15.3	10.0	10.1	3.791667	9.255555	0	9.977011	10.06096	96126.76	58661.97	0	3952.161	995.3554	96126.0	1.3707	1.17	0.029	
67342	1.7647	1.3333	3.2727	1.5455	6.2553	1.2632	7.2821	1.5556	6.7778	1.7143	1.1	2.8	6.5	6.8	5.5	0	2.818182	5.518477	6.837607	5.492063	0	17861.72	8744.066	2708.559	543.8883	17862.0	0.9357	1.71	0.043	
67371	7.7333	18	13.471	25.0	20.900	31.778	11.782	19.200	11.091	15.39	23.7	37.4	50.7	29.1	25.1	23.73333	0	0	0	0	601690.1	0	0	0	0	601690.0	1.8449	1.00	0.025	Green - High End Correction
34573	1.8444	1	4.1905	1	3.9375	1.1429	6.8100	1.2222	5.4237	1.3621	0.9	3.2	3.1	6.0	4.8	0	3.190476	3.080357	6.033033	4.785798	0	20221.33	4880.848	2389.846	473.9457	20221.0	0.9677	1.65	0.041	
37610	1.0667	1.5	1.1111	2.8889	1.1852	2.4667	1.3846	15.75	1.3519	18.176	0.6	2.0	1.7	15.1	17.5	0	0	0	15.13462	0	0	0	0	5995.226	0	5995.0	0.6534	2.45	0.061	
41420	2.2308	3	4.6125	7.0889	11.167	16.188	10	20.636	10.091	19.384	3.2	9.7	25.4	28.6	27.5	3.230769	9.701389	0	0	0	81906.82	81487.68	0	0	0	81906.0	1.3294	1.20	0.030	
41475	1	1	1.1176	1.2222	1.2917	1.4	1.5	1.4167	1.651	1.6471	0.0	0.3	0.7	0.9	1.2	0	0	0	0	0	0	0	0	0	0	0.0	#NUM!	12.00	0.300	Yellow -Low End Correction
41541	3	1.1429	7.9231	1.375	13.063	1.3529	12.824	1.5833	12.911	2.359	2.1	7.3	12.4	12.4	13.3	2.142857	7.298077	0	0	0	54325.96	46255.42	0	0	0	54325.0	1.2232	1.31	0.033	
44307	1.2308	1.1	1.7333	1.1667	5.8846	1.2917	6.25	1.4545	10.564	1.4545	0.3	0.9	6.2	7.7	10.0	0	0	6.176282	7.704545	10.01818	0	0	8201.855	3051.977	992.1175	8201.0	0.7344	2.18	0.054	
44327	1.125	3.8889	1.4211	11	3.0682	22.087	2.8231	19.944	3.4091	14.953	3.0	10.4	23.2	20.9	16.4	3.013889	10.42105	0	0	0	76408.45	66048.92	0	0	0	76408.0	1.3114	1.22	0.031	
45046	1.6875	2	3.8889	4.6667	5.0667	5.9412	6.5641	18.467	6.606	18.029	1.7	6.6	9.0	25.0	22.7	0	6.555555	9.007843	0	0	0	41549.29	14272.99	0	0	41549.0	1.1539	1.39	0.035	
45078	1.0714	1.1111	1.1333	1.2	2.3478	2.378	1.675	1.0667	5	3.0303	0.2	0.3	2.7	1.7	6.0	0	0	2.722628	0	6.030303	0	0	4314.337	0	597.1911	4314.0	0.5684	2.82	0.070	
45080	1.1429	1.4444	1.125	2.5	1.3182	2.1875	1.29	10.167	1.5106	15.25	0.6	1.6	1.5	9.4	14.8	0	0	0	9.446667	0	0	0	0	3742.078	0	3742.1	0.5316	3.01	0.075	
47830	1.8	2	3.4706	4.1	3.4	3.0625	10.807	17.5	10.89	17.005	1.8	5.6	4.5	26.4	26.8	0	5.570588	4.4625	0	0	0	35306.54	7070.663	0	0	35307.0	1.1118	1.44	0.036	
47833	1.125	1.1111	1.3	1.1818	10.738	1.2857	2.6098	1.2778	3.7845	1.3125	0.2	0.5	10.0	1.9	3.1	0	0	10.02156	0	3.066991	0	0	15879.24	0	303.7293	15879.0	0.9053	1.77	0.044	
47857	2	1.1818	3.6667	1.2308	8.3416	1.1538	6.0784	1.3158	3.3187	1	1.2	2.9	7.5	5.4	2.3	0	2.897436	7.495309	5.394221	2.316725	0	18364.03	11876.37	2136.795	229.6273	18364.0	0.9428	1.70	0.042	
47864	18.5	4	17.583	8.9091	30.333	15.278	14.035	19.385	18.933	19.368	20.5	24.6	43.6	31.5	38.3	20.5	0	0	0	0	519718.3	0	0	0	0	519718.0	1.8070	1.00	0.025	Green - High End Correction
47877	3.5625	1.2222	8.0526	1.4545	4.0811	1.1429	8.3448	2.0625	8.3021	2.2069	2.8	7.5	3.2	9.4	8.5	2.784722	7.507177	3.223938	9.407328	8.50898	70598.59	47580.7	5108.353	3726.494	842.6586	70598.0	1.2909	1.24	0.031	
47878	1.2143	5.75	1.6333	16	2.9455	33.875	4.05	22.667	11.543	21.622	5.0	15.5	34.4	24.7	31.2	4.964286	0	0	0	0	125855.1	0	0	0	0	125855.0	1.4404	1.11	0.028	
47880	2.25	1.125	5.0526	1.1	19.486	2.5294	9.6207	1.5385	8.6581	1.8318	1.4	4.2	20.0	9.2	8.5	0	4.152631	0	9.159151	8.490958	0	26319.49	0	3628.185	840.674	26319.0	1.0359	1.54	0.039	
47899	1.3333	1	2.7647	1	2.9333	1.3125	9.6957	1.3077	11.743	1.8718	0.3	1.8	2.2	9.0	11.6	0	0	2.245833	9.003344	11.61373	0	0	3558.539	3566.466	1150.127	3558.0	0.5186	3.09	0.077	
47905	2.2	1.1429	4.1579	1.2222	11.946	1.4286	9.5357	1.7273	8.0636	2.4945	1.3	3.4	11.4	9.3	8.5	0	3.380117	11.37452	9.262987	8.524312	0	21423.28	18023	3669.317	844.1771	21423.0	0.9827	1.63	0.041	
47906	15.385	1.25	17.333	1.3636	17.973	3.4333	10.292	1.9	7.5393	2.1075	14.6	16.7	19.4	10.2	7.6	14.63481	0	0	10.19167	7.646852	371018.4	0	0	4037.192	757.2807	371018.0	1.7199	1.00	0.025	Green - High End Correction
47909	1.0769	2.5	1.0714	6.125	1.3	9	1.3529	21.818	2.32	30.238	1.6	5.2	8.3	21.2	30.6	0	5.196429	8.3	0	0	0	32935.11	13151.41	0	0	32935.0	1.0938	1.46	0.037	
47917	1	4.1111	1.0625	10.5	1.2288	12.55	1.5714	20.571	2.7222	18.486	3.1	9.6	11.8	20.1	19.2	3.111111	9.5625	11.77857	0	0	78873.24	60607.39	18663.23	0	0	78873.0	1.3196	1.21	0.030	
47920	1	1.1111	0.9412	1.3	1.122	5.3333	1.129	3.0769	1.1681	7.1034	0.1	0.2	4.5	2.2	6.3	0	0	4.455285	2.205955	6.291567	0	0	7059.43	873.838	823.0645	7059.0	0.6957	2.30	0.057	
47924	1	1.6667	1	3.1818	1.3143	35.824	1.129	14.607	1.4198	18.653	0.7	2.2	35.1	13.8	18.0	0	2.181818	0	0	0	0	13828.43	0	0	0	13828.0	0.8595	1.84	0.046	
47941	3.2667	4.5556	8.25	12.7	13.706	15.357	12.917	20.067	16.957	23.156	5.8	18.9	27.1	31.0	38.1	5.822222	0	0	0	0	147605.8	0	0	0	0	147606.0	1.4816	1.08	0.027	
47960	1.1875	1.2222	1.5556	1.2727	4.5	1.3529	5.4848	1.5556	9.2619	1.9787	0.4	0.6	3.9	5.0	9.2	0	0	3.852941	5.040404	9.240628	0	0	6105.012	1996.639	915.115	6105.0	0.6581	2.43	0.061	
47963	1.2143	1.2222	1.7333	1.5	5.75	1.6364	6.8571	1.8333	23.412	3.05	0.4	1.2	5.4	6.7	24.5	0	0	5.386364	6.690476	0	0	0	8534.731	2650.277	0	8534.0	0.7447	2.15	0.054	
47972	1.0667	12.222	1.0583	28.2	1.28	52.385	1.2083	24.817	1.5036	33.577	11.3	25.3	51.7	24.1	33.1	11.28889	0	0	0	0	286197.2	0	0	0	0	286197.0	1.6528	1.00	0.025	Green - High End Correction
47981	1.1875	1.1111	1.2778	1.1818	1.5161	1.2273	1.3704	1.3125	1.6882	1.3958	0.3	0.5	0.7	0.7	1.1	0	0	0	0	0	0	0	0	0	0	1177.0	0.2326	12.00	0.300	
53478	15.6	1.25	16.706	1.6667	32.478	2.2907	13.630	4.3333	22.711	12.056	14.9	16.4	32.7	16.0	32.8	14.85	0	0	0	0	376478.9	0	0	0	0	376479.0	1.7237	1.00	0.025	Green - High End Correction
53528	1.0667	1.25	1.2941	1.4444	11.13	1.6154	3.5714	1.9167	12.78	2.1613	0.3	0.7	10.7	3.5	12.9	0	0	10.74582	3.488095	0	0	0	17026.63	1381.728	0	17027.0	0.9233	1.73	0.043	
53529	3.2143	1.5	7.8125	2.7778	21.250	11.250	11.429	13.091	11.356	23.139	2.7	8.6	30.5	22.5	32.6	2.714286	8.590278	0	0	0	68812.87	54445.42	0	0	0	68812.0	1.2843	1.25	0.031	
53537	1.2857	1.125	2.125	1.2222	27.5	2.3846	10.36	1.7273	14.633	2.9375	0.4	1.3	27.9	10.1	15.6	0	0	0	10.07727	0	0	0	0	3991.878	0	3991.0	0.5483	2.92	0.073	
53539	1.1667	1.1429	1.8231	1.25	12.333	1.3529	9.125	1.5833	15.2	2	0.3	1.2	11.7	8.7	15.2	0	0	11.68827	8.708333	0	0	0	18516.98	3449.604	0	18517.0	0.9450	1.69	0.042	
47986	1	1.125	1.0625	1.1111	1.96	2.8667	1.2913	1.6154	2.2128	2.0677	0.1	0.2	2.8	1.0	3.2	0	0	2.826667	0	3.160508	0	0	4478.873	0	314.9711	4478.0	0.5780	2.77	0.069	
47987	2	1.3333	4	1.7778	8.1905	1.6923	13.687	4.4545	41.813	18.571	1.3	3.8	7.9	16.1	58.4	0	3.777778	7.882784	0	0	0	23943.86	12490.33	0	0	23944.0	1.0114	1.58	0.040	
47994	7.4286	16.8	13.5	27.909	29.958	36.632	12.64	18	17.723	18.932	22.2	39.4	64.6	28.6	34.7	22.22857	0	0	0	0	563541.2	0	0	0	0	563541.0	1.8279	1.00	0.025	Green - High End Correction
47997	0.8889	10.556	1	23.5	1.6338	38.9	2.1538	22.571	5.9868	30.963	9.4	22.5	38.5	22.7	34.8	9.444444	0	0	0	0	239436.6	0	0	0	0	239436.0	1.6067	1.00	0.025	Green - High End Correction
48018	1.2353	1	2	1.0769	17.618	1.351	7.1944	1.3182	7.8571	1.625	0.2	1.1	17.0	6.5	7.5	0	0	0	6.512627	7.482143	0	0	0	2579.826	740.9693	2579.8	0.4355	3.67	0.092	
48037	1.125	1.1111	1.3158	1.3	2.0947	1.3846	3.2947	1.6429	5.9823	2.2047	0.2	0.6	2.3	2.9	6.2	0	0	2.279352	2.907563	6.187036	0	0	3611.65	1151.764	612.7126	3611.6	0.5224	3.06	0.077	
53415	1.5	1.4	2.6842	1.4545	7.1808	1.8571	6.1053	1.8462	3.8563	2.6178	0.9	2.1	7.0	6.0	4.5	0	2.138756	7.037994	5.951417	4.474129	0	13555.6	11151.75	2357.516	443.0805	13555.0	0.8643	1.85	0.046	
53416	1.3889	28.778	2	25.636	2.5556	42.188	4.8889	20.286	3.8183	23.385	28.2	25.8	42.7	23.0	25.0	28.16667	0	0	0	0	714084.5	0	0	0	0	714085.0	1.8891	1.00	0.025	Green - High End Correction
53433	1.1875	1	1.2778	1.2	4.8	1.1875	1.6384	1.4615	2.9015	1.8211	0.2	0.5	4.0	1.1	2.9	0	0	3.9875	0	2.882591	0	0	6318.222	0	285.4679	6318.0	0.6670	2.40	0.060	
53444	10.733	1.25	13.471	1.4444	28.622	2.1667	12	2.4545	16.113	7	10.0	14.9	28.9	12.5	21.1	9.983334	0	0	0	0	253098.6	0	0	0	0	253098.0	1.6210	1.00	0.025	Green - High End Correction
53459	2.125	1.1111	4.6842	1.3	22.316	1.4375	10.154	1.8288	11.318	2.4634	1.2	4.0	22.0	10.1	11.8	0	3.98421	0	10.08242	11.77987	0	25252.04	0	3993.915	1166.58	25252.0	1.0252	1.56	0.039	
53522	1.2	1.75	1.7647	3.5556	5.0345	8.8889	7.1818	14.769	11.441	17.638	1.0	3.3	11.9	20.0	27.1	0	3.320281	11.92337	0	0	0	21043.91	18892.67	0	0	21044.0	0.9781	1.64	0.041	
53544	1.1667	1	1.6023	1.2917	2.8696	1.7447	6.75	4.1429	12.18	5.6111	0.2	1.0	2.6	8.9	15.8	0	0	2.614246	8.892857	0	0	0	4142.291	3522.699	0	4142.3	0.5579	2.87	0.072	
54762	2.4375	1.375	5.1579	1.5658	17.947	1.7273	6.4657	1.3846	6.901	2.3182	1.8	4.7	17.7	5.9	7.2	0	4.71345	0	5.87033	7.219172	0	29873.98	0	2325.395	714.9268	29874.0	1.0686	1.50	0.037	
54777	0.9444	17.167	1.0435	17.5	1.22	18.735	1.8837	10	3.0349	7.9424	16.1	16.5	18.0	9.9	9.0	16.11111	0	0	9.883721	8.977243	408450.7	0	0	3915.206	889.0316	408451.0	1.7447	1.00	0.025	Green - High End Correction
54915	1	12.583	1.0478	16.188	1.1053	13.727	1.8667	9.9667	3.019	7.697	11.6	15.2	12.8	9.6	8.7	11.58333	0	0	9.633333	8.716017	293662	0	0	3816.021	863.1619	293662.0	1.6594	1.00	0.025	Green - High End Correction
66385	1	1	1.0526	1.0909	1.0256	1.0526	1.1	1.2667	1.2255	1.5909	0.0	0.1	0.1	0.4	0.6	0	0	0	0	0	0	0	0	0	0	80.0	-0.4624	12.00	0.300	
66334	1	1.2222	1.1111	1.5	1.1379	2.1111	1.3478	4.7692	1.5645	12.45	0.2	0.6	1.2	4.1	12.0	0	0	0	4.117057	12.01452	0	0	0	1630.876	1189.818	1630.0	0.3168	5.05	0.126	
66344	1.6875	1.1111	3.5556	1.2	33.909	2.7143	8.9048	1.75	20.489	4.6364	0.8	2.8	34.6	9.7	23.1	0	2.755556	0	9.654762	0	0	17464.79	0	3824.51	0	17465.0	0.9299	1.72	0.043	
66343	1	1.1	1	1.1818	1.1053	1.1154	1.075	1.2353	1.1222	1.1067	0.1	0.2	0.2	0.3	0.3	0	0	0	0	0	0	0	0	0	0		#NUM!	12.00	0.300	
67309	1.3333	2.25	2	4.7	7.4348	10.058	8.4286	17.429	18.016	16.167	1.6	4.7	15.5	24.9	32.2	0	4.7	0	0	0	0	29788.73	0	0	0	29789.0	1.0679	1.50	0.037	
67340	1.7857	3.5	3.7333	8.2	4.5	5.7727	8.7308	14.444	11.276	11.429	3.3	9.9	8.3	22.2	20.7	3.285714	9.933333	8.272727	0	0	83299.8	62957.75	13108.19	0	0	83300.0	1.3337	1.20	0.030	
67356	1.8125	2.25	3.8111	4.6	16.278	33.889	8.0333	16.8	8.1875	16.637	2.1	6.4	50.2	23.8	24.0	2.0625	6.411111	0	0	0	52288.73	40633.8	0	0	0	52289.0	1.2133	1.32	0.033	
67363	8.0638	1.6667	10.042	3.6	12.038	8.2105	7.4595	10.824	7.3363	12.646	8.5	11.6	19.2	16.3	18.0	8.530303	11.64167	0	0	0	216261.2	73785.21	0	0	0	216261.0	1.5603	1.01	0.025	
67382	1.5625	1.1111	3.0556	1.0909	14.679	1.2778	7.875	1.2667	11.319	1.4815	0.7	2.1	14.0	7.1	10.8	0	2.146465	0	7.141667	10.78038	0	13604.35	0	2829.005	1067.599	13604.0	0.8653	1.85	0.046	
47891	1.0667	1.125	1.0588	1.1	1.1471	1.0435	1.125	1.2	1.3222	1.1579	0.2	0.2	0.2	0.3	0.5	0	0	0	0	0	0	0	0	0	0		#NUM!	12.00	0.300	
53531	1.3125	1	2.1053	1.0909	18.065	1.2	7.5769	1.5	9.4874	1.4091	0.3	1.2	15.3	7.1	8.9	0	0	0	7.076923	8.876462	0	0	0	2803.359	879.053	2803.0	0.4569	3.50	0.088	
41064	1.0625	1.6889	1.0478	3.4167	1.1471	3.1389	1.2703	8.3182	1.2917	5.5768	1.0	2.5	2.3	7.6	4.9	0	2.464286	2.285948	7.588452	4.868309	0	15618.71	3622.1	3005.989	482.1169	15619.0	0.9010	1.78	0.044	
41466	1.0667	4.125	1.0526	10.8	1.1455	18.75	1.2581	18.5	1.7517	23.288	3.2	9.9	15.9	17.8	23.0	3.191667	9.852632	0	0	0	80915.49	62446.26	0	0	0	80915.0	1.3262	1.21	0.030	
41452	0.8286	1	1	1.2857	1.1724	1.4615	1.0455	1.625	1.1852	2.1733	-0.1	0.3	0.6	0.7	1.4	0	0	0	0	0	0	0	0	0	0	1004.0	0.1915	12.00	0.300	
47901	0.8824	0.8889	0.8571	1	1.3455	1.2632	1.0611	1.2	1.4143	1.4376	-0.2	-0.1	0.6	0.3	0.9	0	0	0	0	0	0	0	0	0	0	964.0	0.1810	12.00	0.300	

Figure 28

47933	0.9375	16.556	1.1111	21.8	1.3214	43.487	2.375	19.167	7.2529	27.071	15.5	20.7	42.8	19.5	32.4	15.49306	0	0	0	0	392781.7	0	0	0	0	392782.0	1.7346	1.00	0.025	Green - High End Correction
46094	2.0667	1.125	4.6235	1.2222	23.68	1.3571	10.281	1.5	14.291	2.25	1.2	4.0	23.0	9.8	14.5	0	4.045751	0	9.76087	0	0	25642.09	0	3868.542	0	25642.0	1.0291	1.55	0.039	
46095	1.1765	1.1	1.8421	1.0760	4.6944	1.125	5.7333	1.2105	7.0062	1.473	0.3	0.9	3.8	4.9	7.1	0	0	3.619445	4.943659	7.06122	0	0	6051.937	1958.395	701.2652	6051.9	0.6559	2.44	0.061	
83412	1	1.2857	1	1.5	1.067	1.5385	1.05	3.5	1.6316	17.75	0.3	0.5	0.6	2.5	17.4	0	0	0	2.55	17.38158	0	0	0	1010.123	1721.327	1010.0	0.1930	8.29	0.207	
83475	1	1	1	1	1.1481	1.2667	1.5714	1.1818	4.9531	1.7	0.0	0.0	0.4	0.8	4.7	0	0	0	0	4.653125	0	0	0	0	460.6068	460.0	-0.0103	12.00	0.300	
83503	1.2	1.375	1.2353	1.5556	1.2	1.4815	1.7143	2.6667	4.3654	8.625	0.6	0.8	0.7	2.4	8.0	0	0	0	2.380952	7.990385	0	0	0	943.159	791.3013	943.0	0.1753	9.13	0.228	
83540	7.6235	13.676	8.7	18.2	6.0068	8.1333	7.675	13.308	13.708	21.725	19.7	22.9	10.2	19.2	33.4	19.69853	0	10.23011	0	0	499399.3	0	16209.68	0	0	499399.0	1.7967	1.00	0.025	Green - High End Correction
83580	0.875	3	0.8412	8	1.1905	10.222	1.2532	13.167	2.2647	17.686	1.9	6.9	9.4	12.4	18.0	0	6.941176	9.412699	0	0	0	43993.37	14914.49	0	0	47535.0	1.1887	1.35	0.034	
66334	1.2667	1.125	2.2353	1.5	23.692	2	9.1053	2.2	14.64	2.0667	0.4	1.7	23.7	9.3	15.3	0	0	0	9.305263	15.30667	0	0	0	3686.064	1515.845	3686.0	0.5277	3.03	0.076	
47967	2	4	3.75	8.4444	20.869	25.5	7.675	11.75	10.758	10.305	4.0	11.2	44.4	17.6	19.1	4	11.19444	0	0	0	101408.5	70950.71	0	0	0	101408.0	1.3846	1.18	0.029	
46083	1.2632	1.1111	1.7308	1.1018	2.3908	1.069	3.6087	1.1579	2.8067	1.3043	0.4	0.9	1.5	2.8	2.2	0	0	0	2.824561	2.241057	0	0	0	1118.884	221.9356	1118.9	0.2195	7.29	0.182	
83533	11.087	1	11.045	1.2222	16.133	4.5	7.0857	2	6.6777	4.3143	10.7	10.3	17.6	7.1	9.0	10.86667	10.26768	0	7.085714	8.991971	270422.5	65076.83	0	2806.841	890.4901	270423.0	1.6381	1.00	0.025	Green - High End Correction
37979	3.2	1	8.1176	1.2222	17.242	1.3571	8.7917	1.5	10.737	2.8	2.2	7.3	16.6	8.3	11.2	2.2	7.339869	0	8.291667	11.23684	55774.85	46520.3	0	3284.551	1112.803	55774.0	1.2300	1.30	0.033	
41363	1.0667	1	1.1765	1.1667	3.6333	2.913	2.7917	2.4	8.675	6.0725	0.1	0.3	4.7	3.2	13.9	0	0	4.746377	3.191667	0	0	0	7520.667	1264.305	0	7520.0	0.7120	2.25	0.056	
41548	1.4375	1.375	1.65	1.9091	1.6723	1.6519	4.0606	5.6333	5.1355	6.2727	0.8	1.8	1.7	7.9	11.4	0	0	0	7.893939	11.40821	0	0	0	3127.001	1129.774	3127.0	0.4852	3.30	0.082	
47852	8.2351	1.125	7.9615	1.6667	8.6533	2.6947	4.3023	4.1638	4.0084	10.687	7.4	7.6	10.7	6.5	12.7	7.363095	7.628205	10.74807	6.456172	0	186670	48347.78	17030.39	2557.462	0	186670.0	1.5423	1.04	0.026	
47885	1.0667	1.1429	1.2941	1.25	3.6065	1.4286	3.067	1.5	8.5132	1.9143	0.2	0.5	3.2	2.6	8.4	0	0	3.235023	2.586957	8.427444	0	0	5125.917	1024.763	834.584	5125.0	0.6129	2.61	0.065	
47832	1.5882	1.125	2.0610	1.2222	13.479	1.2941	5.9677	1.4945	6.5287	1.814	0.7	1.9	12.8	5.4	6.4	0	0	0	5.422287	6.402663	0	0	0	2147.913	634.0666	2147.0	0.3880	4.12	0.103	
47815	4.5714	1	8.6296	1.1	7.1263	1.5	3.6228	1.375	3.0478	2.1515	3.6	5.7	6.6	3.0	3.2	3.571429	5.72963	6.626316	2.997642	3.199297	90543.26	36314.55	10499.44	1187.446	318.8318	90543.0	1.3553	1.18	0.030	
47925	10.294	1.75	9.8571	4	13.365	23.5	6.0958	11.083	6.4259	14.96	10.0	11.9	40.9	15.2	18.0	10.04412	11.85714	0	0	0	254639.6	75150.9	0	0	0	254640.0	1.6226	1.00	0.025	Green - High End Correction
47980	12.212	1.375	12.789	2.2222	18.116	0.75	8.0714	8.7273	7.4281	18.641	11.7	13.0	20.9	14.8	22.1	11.6875	0	0	0	0	296302.8	0	0	0	0	296303.0	1.6617	1.00	0.025	Green - High End Correction
37319	8.75	1.2857	11.618	1.75	20.682	1.625	10.294	8.4545	15.325	12.917	8.0	11.4	20.3	14.7	26.2	8.035714	11.36539	0	0	0	203722.3	72034.13	0	0	0	203722.0	1.5649	1.02	0.026	
29043	1.6923	1.1429	2.7333	1.25	8.5405	1.3	7.9565	1.5385	7.6077	1.8664	0.8	2.0	1.8	7.5	7.7	0	0	0	7.494983	7.694056	0	0	0	2968.963	761.9554	2969.0	0.4718	3.39	0.085	
83443	11.667	1.2857	14.059	1.875	18.061	9.0760	9.2580	8.5	8.6477	21.2	11.0	13.9	25.6	13.8	27.8	10.95238	0	0	0	0	277668	0	0	0	0	277666.0	1.6450	1.00	0.025	Green - High End Correction
41429	1.0833	1.1429	1.2308	1.2857	1.84	1.4815	1.6471	1.5556	3.7143	2.4286	0.2	0.5	1.3	1.2	4.1	0	0	0	0	4.142857	0	0	0	0	410.2741	410.0	-0.0400	12.00	0.300	
47837	1.0714	1.1429	1.2667	1.375	1.4783	1.4167	3.4737	1.8	12.462	2.7619	0.2	0.6	0.9	3.3	13.2	0	0	0	3.273684	0	0	0	0	1296.794	0	1296.0	0.2575	6.21	0.155	
41184	1.1429	2.125	1.0588	5	1.75	36.178	1.2009	16.184	1.6421	16.222	1.3	4.1	35.9	15.4	16.1	0	4.058824	0	0	0	0	25724.94	0	0	0	32142.0	1.0875	1.47	0.037	
66302	1.1538	1.1429	1.3125	1.25	15.803	1.5	3.2913	1.7	8.9412	2.75	0.3	0.6	15.4	3.1	9.7	0	0	0	3.091304	9.691176	0	0	0	1224.848	959.7336	1224.0	0.2427	6.59	0.165	
83464	1.7692	1.1429	3.7333	1.25	3.4444	1.6333	8.9	2.5568	13.537	10.381	0.9	3.0	3.0	9.2	23.9	0	2.983333	3.027778	9.155556	0	0	18908.45	4797.535	3626.761	0	4797.0	0.5958	2.69	0.067	
83478	10.231	1.2857	12.5	1.2222	24.333	1.3684	10.844	1.6667	16.308	2.5385	9.5	12.0	23.7	10.6	15.8	9.516483	0	0	10.61111	0	241263	0	0	4203.345	0	241263.0	1.6086	1.00	0.025	Green - High End Correction
47919	2.4	1.25	7.2778	1.3	17.514	1.2917	8.3077	1.6471	10.183	3.25	2.7	6.6	16.8	8.0	11.4	2.65	6.577778	0	7.954751	11.43519	67183.1	41690.14	0	3151.09	1132.446	67183.0	1.2781	1.25	0.031	

Figure 29a

107 Plex	40 minute read					Optimized Primer ConcentrationCalls Made at 40 min read
101 SNPs	Fam FOZ	Red FOZ	Dye Call	Final Call	Genomic Call	94
47887	4.2832	5.6	HET	HET	HET	1
53488	16.697	2.578	FAM	C	FAM	1
67329	8.6172	1.5	FAM	C	FAM	1
67342	4.4773	1.1818	FAM	C	FAM	1
67371	18.233	32.278	HET	HET	HET	1
34573	4.371	1	FAM	C	FAM	1
37810	1.2727	8.2832	RED	G	RED	1
41420	13.608	19.65	HET	HET	HET	1
41475	1.3333	1.3333	HET	HET	HET	1
41541	12.963	1.4286	FAM	G	FAM	1
44907	4.24	1.2	FAM	C	FAM	1
44927	1.7727	19.048	RED	T	RED	1
45048	3.6923	4.6667	HET	HET	HET	1
45078	1.8696	1.7222	HET	HET	HET	1
45080	1.4348	11.111	RED	T	RED	1
47850	3.9259	3.5789	HET	HET	HET	1
47855	1.9268	1.35	43	43	FAM	0
47857	9.3951	1.1429	FAM	C	FAM	1
47868	29.519	3.8182	FAM	C	FAM	1
47877	14.171	1.32	FAM	C	FAM	1
47878	1.4074	5.75	RED	A	RED	1
47880	5.6571	1.4091	FAM	C	FAM	1
47899	11.484	1.5	FAM	G	FAM	1
47905	6.1591	1.1905	FAM	C	FAM	1
47906	20.053	1.3836	FAM	G	FAM	1
47909	1.7222	22.214	RED	T	RED	1
47917	2.4231	28.706	RED	A	RED	1
47920	1.3077	2.2353	RED	A	RED	1
47924	1.2667	6.8889	RED	A	RED	1
47941	8.2174	9	HET	HET	HET	1
47960	5.6697	2.0526	FAM	G	FAM	1
47963	6.9091	1.6667	FAM	G	FAM	1
47972	1.56	40.941	RED	G	RED	1
47981	2.6316	1.2593	FAM	G	FAM	1
53478	26.963	1.65	FAM	C	FAM	1
53528	1.8571	1.4	HET	HET	FAM	0
53529	17.103	10.19	HET	HET	HET	1
53557	7.6364	1.3182	FAM	G	FAM	1
53530	5.3846	1.7	FAM	C	FAM	1
47986	1.4	1.6875	HET	HET	HET	1
47987	12	1.7333	FAM	G	FAM	1
47994	23.583	31.714	HET	HET	HET	1
47997	1.4615	8.3125	RED	A	RED	1
48088	6.4	1.2	FAM	G	FAM	1
48097	3.7222	1.65	FAM	G	FAM	1
53415	3.5904	1.8471	FAM	G	FAM	1
53418	1.3176	38.85	RED	A	RED	1
53433	1.2903	1.2273	39	39	FAM	0
53444	18.152	1.8667	FAM	G	FAM	1
53459	5.675	1.4167	FAM	G	FAM	1
53522	2.7273	4.6606	HET	HET	HET	1
53544	8.8774	4.8955	HET	HET	HET	1
54762	5.1844	1.3571	FAM	C	FAM	1
54777	1.1136	11.912	RED	A	RED	1
54915	1.2143	10.368	RED	A	RED	1
66865	1.05	1.1429	59	59	FAM	0
66934	1.1389	7.9565	RED	A	RED	1
66984	7.7931	2.05	FAM	G	FAM	1
66985	1.1	1.2	39	39	FAM	0
67309	2.9444	6.4483	HET	HET	HET	1
67340	4.0769	5.8095	HET	HET	HET	1
67358	8.0303	6.55	HET	HET	HET	1
67363	1.8298	1.1765	FAM	G	FAM	1
67332	18.387	1.2083	FAM	G	FAM	1
87801	12.333	1.3704	FAM	C	FAM	1
53551	6.913	1.129	FAM	C	FAM	1
41084	1.4545	5.3864	RED	A	RED	1
41408	1.1688	20.522	RED	A	RED	1
41452	1.3158	1.6111	HET	HET	FAM	0
47901	1.7692	1.3638	HET	HET	FAM	0
47932	1.3214	13.765	RED	A	RED	1
48094	5.8519	1.3125	FAM	C	FAM	1
48095	12.122	1.1768	FAM	C	FAM	1
53412	1.5172	11.222	RED	A	RED	1
53475	10.412	1.3333	FAM	G	FAM	1
53503	3.9375	3.9444	HET	HET	HET	1
53540	17.595	18.727	HET	HET	HET	1
55560	1.4333	10.923	RED	A	RED	1
66938	9.3333	1.819	FAM	C	FAM	1
47867	4.9259	7.4762	HET	HET	HET	1
48083	4.3	1.1389	FAM	G	FAM	1
53553	14.34	1.4375	FAM	C	FAM	1
37879	7.5429	1.6	FAM	T	FAM	1
41363	2.2333	2.1071	HET	HET	HET	1
41846	8.2105	9.2059	HET	HET	HET	1
47852	10.244	1.3913	FAM	C	17	1
47885	2.3421	1.2632	FAM	T	FAM	1
47892	8.9508	1.2083	FAM	G	FAM	1
47915	8.6667	1.1788	FAM	G	17	1
47925	16.898	3.6364	FAM	G	FAM	1
47980	18.963	1.6364	FAM	G	FAM	1
37319	19.478	1.6471	FAM	C	FAM	1
29043	10.818	1.4211	FAM	G	FAM	1
53445	12.615	1.8	FAM	C	17	1
41409	12.154	1.8	FAM	C	FAM	1
47897	5.0417	1.6667	FAM	G	FAM	1
41164	1.5357	7.2	RED	T	RED	1
66902	13.875	1.7059	FAM	G	FAM	1
53464	21.75	2.5294	FAM	C	FAM	1
53470	14.467	1.4231	FAM	C	17	1
47810	16.159	1.2759	FAM	C	FAM	1

CNR; "Call Not Reproduced", NS; "New Signal"

Figure 29b

Compare 10 min Read of Balanced Mix (MAP Invader 092701) v. Optimized Mix (MAP Invader 100301)

Primer Set Use	10 min Calls with Balanced Mix					10 min Calls w/ Optimized Mix						Bal.	Opt.
107-plex (101)	Fam FOZ	Red FOZ	FOZ RATIO	Dye Call	Final Call	Fam FOZ	Red FOZ	FOZ RATIO	Dye Call	Final Call	Genomic Call	50	71
47887	1.25	1.375	0.6668	038	38	1.4375	1.4167	1.0600	HET	HET	HET	0	1
53488	1.1176	1.1	1.1764	39	39	13.5	1.0667	187.4900	FAM	C	FAM	0	1
67329	4.6667	1.125	29.3330	FAM	C	1.65	1	64.9990	FAM	C	FAM	1	1
67342	1.7647	1.3333	2.2941	HET	HET	1.44	0.9333	44.0000	17	17	FAM	0	0
67371	7.7333	18	0.3960	53	53	3.6316	6.6385	0.4751	HET	HET	HET	0	1
34573	1.9444	1	94.4440	FAM	C	1.5	0.8235	50.0000	17	17	FAM	1	0
37810	1.0667	1.5	0.1333	28	28	1.0833	1.7857	0.1060	RED	G	RED	0	1
41420	2.2308	3	0.6153	HET	HET	2.9545	3.3571	0.8292	HET	HET	HET	1	1
41475	1	1	1.0000	39	39	1.0909	1.0667	1.3636	39	39	HET	0	0
41541	3	1.1429	13.9990	FAM	G	2.7368	1.0714	24.3150	FAM	G	FAM	1	1
44807	1.2308	1.1	2.3076	39	39	1.4444	1	44.4440	17	17	FAM	0	0
44927	1.125	3.8889	0.0432	RED	T	1.2727	3.4	0.1136	RED	T	RED	1	1
45046	1.6875	2	0.6876	HET	HET	1.4091	1.4615	0.8863	HET	HET	HET	1	1
45076	1.0714	1.1111	0.6428	39	39	1.1053	1	10.5260	19	19	HET	0	0
45080	1.1429	1.4444	0.3214	68	68	1.1111	1.9231	0.1203	RED	T	RED	0	1
47850	1.8	2	0.7999	HET	HET	1.4737	1.3846	1.2315	HET	HET	HET	1	1
47855	1.125	1.1111	1.1249	39	39	1.1905	1.0714	2.6666	49	49	FAM	0	0
47857	2	1.1818	5.5000	FAM	C	6.2667	0.95	526.6600	FAM	C	FAM	1	1
47868	18.5	4	5.8333	FAM	C	20.3	1.1765	109.3600	FAM	C	FAM	1	1
47877	3.5625	1.2222	11.6310	FAM	C	3.36	0.9412	235.9900	FAM	C	FAM	1	1
47878	1.2143	5.75	0.0451	RED	A	1.0952	1.4667	0.2040	28	28	RED	1	0
47880	2.25	1.125	10.0000	FAM	C	1.56	1.0625	8.9599	17	17	FAM	1	0
47899	1.3333	1	33.3330	17	17	2.3478	1.1333	10.1080	FAM	G	FAM	0	1
47905	2.2	1.1429	8.3999	FAM	C	1.7692	1.0667	11.5380	FAM	C	FAM	1	1
47906	15.385	1.25	67.5380	FAM	G	7.6667	1.1053	63.3330	FAM	G	FAM	1	1
47909	1.0769	2.5	0.0512	RED	T	1.26	3.25	0.1111	RED	T	RED	1	1
47917	1	4.1111	0.0032	RED	A	1.4762	4.5385	0.1345	RED	A	RED	1	1
47920	1	1.1111	0.0899	29	29	1.1053	1.1538	0.6842	39	39	RED	0	0
47924	1	1.6667	0.0160	RED	A	1.0909	1.4286	0.2121	28	28	RED	1	0
47941	3.2667	4.5556	0.6375	HET	HET	1.8235	1.75	1.0980	HET	HET	HET	1	1
47960	1.1875	1.2222	0.8437	39	39	1.6316	1.2308	2.7368	41	41	FAM	0	0
47963	1.2143	1.2222	0.9642	39	39	1.875	1	67.5000	FAM	G	FAM	0	1
47972	1.0667	12.222	0.0059	RED	G	1.2222	7.7692	0.0328	RED	G	RED	1	1
47981	1.1875	1.1111	1.6874	39	39	1.2609	1	26.0860	19	19	FAM	0	0
53478	15.6	1.25	58.4000	FAM	C	5.35	1.1333	32.6250	FAM	C	FAM	1	1
53528	1.0667	1.25	0.2666	69	69	1.1364	1	13.6360	19	19	FAM	0	0
53529	3.2143	1.5	4.4285	FAM	G	4.28	1.8824	3.7173	43	43	HET	0	0
53557	1.2857	1.125	2.2857	39	39	1.68	0.9412	67.9990	FAM	G	FAM	0	1
53530	1.1667	1.1429	1.1666	39	39	1.55	1.2143	2.5666	47	47	FAM	0	0
47986	1	1.125	0.0800	29	29	1.1111	1.0769	1.4444	39	39	HET	0	0
47987	2	1.3333	2.9999	43	43	2.1176	1	111.7600	FAM	G	FAM	0	1
47994	7.4286	18.8	0.4068	HET	HET	4.5294	5.9333	0.7154	HET	HET	HET	1	1
47997	0.8889	10.556	0.0010	RED	A	1.1053	1.6154	0.1710	RED	A	RED	1	1
48088	1.2353	1	23.5290	19	19	1.7	1	70.0000	FAM	G	FAM	0	1
48097	1.125	1.1111	1.1249	39	39	1.3636	1.1538	2.3636	37	37	FAM	0	0
53415	1.5	1.4	1.2500	HET	HET	2.1111	1.0769	14.4440	FAM	G	FAM	0	1
53418	1.3889	28.778	0.0139	RED	A	1.12	12.071	0.0108	RED	A	RED	1	1
53433	1.1875	1	18.7500	19	19	1	0.875	1.0000	39	39	FAM	0	0
53444	10.733	1.25	38.9330	FAM	G	3.2727	1.1333	17.0450	FAM	G	FAM	1	1
53459	2.125	1.1111	10.1240	FAM	G	1.6538	1	65.3840	FAM	G	FAM	1	1
53522	1.2	1.75	0.2666	52	52	1.3478	1.4706	0.7391	HET	HET	HET	0	1
53544	1.1667	1	16.6660	19	19	2.2273	1.3636	3.3749	43	43	HET	0	0
54762	2.4375	1.375	3.8333	43	43	1.6	1.0833	7.1999	FAM	C	FAM	0	1
54777	0.9444	17.167	0.0006	RED	A	1	2.9412	0.0051	RED	A	RED	1	1
54915	1	12.583	0.0008	RED	A	1	2.5	0.0066	RED	A	RED	1	1
66865	1	1	1.0000	39	39	0.9524	1	1.0000	39	39	FAM	0	0
66934	1	1.2222	0.0450	29	29	1	1.625	0.0160	RED	A	RED	0	1
66984	1.6875	1.1111	6.1874	FAM	G	1.75	1.1333	5.6250	FAM	G	FAM	1	1
66985	1	1.1	0.0999	29	29	0.9706	1	1.0000	39	39	FAM	0	0
67309	1.3333	2.25	0.2666	53	53	1.25	1.5263	0.4749	38	38	HET	0	0
67340	1.7857	3.5	0.3142	53	53	1.3889	1.5714	0.6805	HET	HET	HET	0	1
67356	1.8125	2.25	0.6500	HET	HET	1.5455	1.5	1.0909	HET	HET	HET	1	1
67363	8.8636	1.6667	11.7950	FAM	G	1.04	1	3.9999	49	49	FAM	1	0
67382	1.5625	1.1111	5.0624	17	17	4.05	1	305.0000	FAM	G	FAM	0	1
67601	1.0667	1.125	0.5333	39	39	2.8261	1.0625	29.2170	FAM	C	FAM	0	1
53551	1.3125	1	31.2500	17	17	1.7778	1.0526	14.7770	FAM	C	FAM	0	1
41064	1.0625	1.8889	0.0703	RED	A	1.1429	1.8182	0.1748	RED	A	RED	1	1
41408	1.0667	4.125	0.0213	RED	A	1.1852	3.6471	0.0899	RED	A	RED	1	1
41452	0.9286	1	1.0000	39	39	1.1304	1.2308	0.5652	39	39	FAM	0	0
47901	0.8824	0.8889	1.0000	39	39	1.2174	1.1538	1.4130	39	39	FAM	0	0
47932	0.9375	16.556	0.0006	RED	A	1.0526	2.3077	0.0402	RED	A	RED	1	1
48094	2.0667	1.125	8.5333	FAM	C	1.5294	1.0909	5.8235	17	17	FAM	1	0
48095	1.1765	1.1	1.7647	39	39	2.8095	1.0667	27.1420	FAM	C	FAM	0	1
53412	1	1.2857	0.0350	29	29	1.2105	2.2308	0.1710	RED	A	RED	0	1
53475	1	1	1.0000	39	39	2.1429	1.0769	14.8570	FAM	G	FAM	0	1
53503	1.2	1.376	0.6333	38	38	1.4762	1.3333	1.4285	HET	HET	HET	0	1
53540	7.8235	13.875	0.5299	HET	HET	3.58	3.8	0.9848	HET	HET	HET	1	1
55560	0.875	3	0.0050	RED	A	1.0417	2.1667	0.0357	RED	A	RED	1	1
66938	1.2667	1.128	2.1333	39	39	2.04	1.2	5.1999	FAM	C	FAM	0	1
47867	2	4	0.3333	53	53	1.4	1.8667	0.4615	HET	HET	HET	0	1
48083	1.2632	1.1111	2.3684	39	39	1.6563	1	65.6250	FAM	G	FAM	0	1
53553	11.667	1	1066.6000	FAM	C	4.8261	1.0833	45.9130	FAM	C	FAM	1	1
37879	3.2	1	220.0000	FAM	T	1.8947	1.0833	10.7360	FAM	T	FAM	1	1
41365	1.0667	1	6.6668	19	19	1.1579	1.1333	1.1842	39	39	HET	0	0
41646	1.4375	1.378	1.1666	HET	HET	2.48	2.2941	1.1436	HET	HET	HET	1	1
47852	8.2381	1.125	57.9040	FAM	C	5.7027	1.0714	65.6370	FAM	C	17	1	1
47885	1.0667	1.1429	0.4666	39	39	1.2381	1.0769	3.0952	49	49	FAM	0	0
47892	1.5882	1.125	4.7058	17	17	2.3548	1	135.4800	FAM	G	FAM	0	1
47915	4.5714	1	357.1400	FAM	G	6.95	1	594.9900	FAM	G	17	1	1
47925	10.294	1.75	12.3920	FAM	G	22.452	1.4286	50.0530	FAM	G	FAM	1	1
47980	12.313	1.375	30.1660	FAM	G	10.036	1.125	72.2850	FAM	G	FAM	1	1
37319	8.75	1.2857	27.1250	FAM	C	3.8824	1.0769	37.4700	FAM	C	FAM	1	1
29043	1.6923	1.1429	4.8461	FAM	G	2.4286	1.0769	18.5710	FAM	G	FAM	1	1
53445	11.667	1.2857	37.3330	FAM	C	2.7727	1.1667	10.6360	FAM	C	17	1	1
41409	1.0833	1.1429	0.5833	39	39	2.2222	1.0833	14.6660	FAM	C	FAM	0	1
47897	1.0714	1.1429	0.4999	39	39	1.3889	1.0909	4.2777	17	17	FAM	0	0
41164	1.1429	2.125	0.1269	RED	T	1.2222	1.6923	0.3209	52	52	RED	1	0
66902	1.1538	1.1429	1.0769	39	39	2.6947	1.0833	22.7360	FAM	G	FAM	0	1
53464	1.7692	1.1429	5.3846	FAM	C	4.6	1.1538	23.4000	FAM	C	FAM	1	1
53470	10.231	1.2857	32.3070	FAM	C	3.25	1.0714	31.5000	FAM	C	17	1	1
47910	3.4	1.25	9.6000	FAM	C	4.25	1.0625	52.0000	FAM	C	FAM	1	1

Figure 30

MAP092701 PCR1_ Optimization

Snp number		10 min		20 min		40 min		80 min		160 min		RED-FOZ + FAM-FOZ -2					Limited RED-FOZ + FAM-FOZ -2					Calculation of Amplification Factor (F)					Selected Value of F	Apparent Primer Concentration	Snp number		Corrected Primer Concentration	Amount of Primers in Optimized Mix
	Primers not added	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	Fam FOZ	Red FOZ	10	20	40	80	160	10	20	40	80	160	10	20	40	80	160				Primers not added		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Plate Layout for Oligo Sets

Figure 31

CYP2D6 PCR amplification:

Primers:

Triplex PCR protocol

Exons 1 & 2 (2036 nt)

2D6L1F1: 5' – CTGGGCTGGGAGCAGCCTC - 3'
2D6L1R1: 5' – CACTCGCTGGCCTGTTTCATGTC – 3'

Exons 3, 4, 5, & 6 (1683 nt)

2D6L2F: 5' – CTGGAATCCGGTGTCGAAGTGG – 3'
2D6L2R2: 5' – CTCGGCCCCTGCACTGTTTC – 3'

Exons 7, 8, & 9 (1754 nt)

2D6L3F: 5' – GAGGCAAGAAGGAGTGTCAGGG – 3'
2D6L3R5B: 5' – AGTCCTGTGGTGAGGTGACGAGG – 3'

Monoplex PCR protocol

CYP2D6 nucleotides 506 – 856 (*10 & *21)

forward (1221-09-01): 5' – ggtagtgaggcaggt –3'
reverse (1221-09-02): 5' – gcttctggtaggggag – 3'

CYP2D6 nucleotides 1335 – 1616 (*11 & *17)

forward (1221-09-03): 5' – aaataggactaggacctgt –3'
reverse (1221-09-04): 5' – gggtcccacggaaat – 3'

CYP2D6 nucleotides 2092 – 2582 (*4, *6 & *37)

forward (1221-09-05): 5' – catggccacgcg –3'
reverse (1221-09-06): 5' – ccggcacctctcg – 3'

CYP2D6 nucleotides 2977 – 3146 (*3 & *33)

forward (1221-09-07): 5' – ccgtcctcctgcat –3'
reverse (1221-09-08): 5' – cactctcaccttctcca – 3'

FIGURE 31 (cont.)

CYP2D6 nucleotides 3294 – 3494 (*2 R296C & *7)

forward (1221-09-09): 5' – gttctgtcccgagtatg –3'
reverse (1221-09-10): 5' – tgcactgtttcccaga – 3'

CYP2D6 nucleotides 3589 – 3918 (*25, *26 & *29)

forward (1221-09-11): 5' – ctgacctcctccaacat –3'
reverse (1221-09-12): 5' – gggctatcaccaggt – 3'

CYP2D6 nucleotides 4316 – 5226 (*2, *27, *31 & *32)

forward (1221-09-13): 5' – ctgacctcctccaacat –3'
reverse (1221-09-15): 5' – gggctatcaccaggt – 3'

FIGURE 32

Protocol for Multiplex PCR Optimization

Designed PCR primers are mixed at equimolar concentrations and used for target pre-amplification.

↓

Amplification factor for each PCR amplicon is measured using Invader SNP assay.

↓

Concentration of PCR primers is adjusted to achieve the unbiased (“Optimized”) amplification.

↓

The adjustment coefficients are stored in the database and may be used for new multiplex PCR pre-amplification without the optimization step.

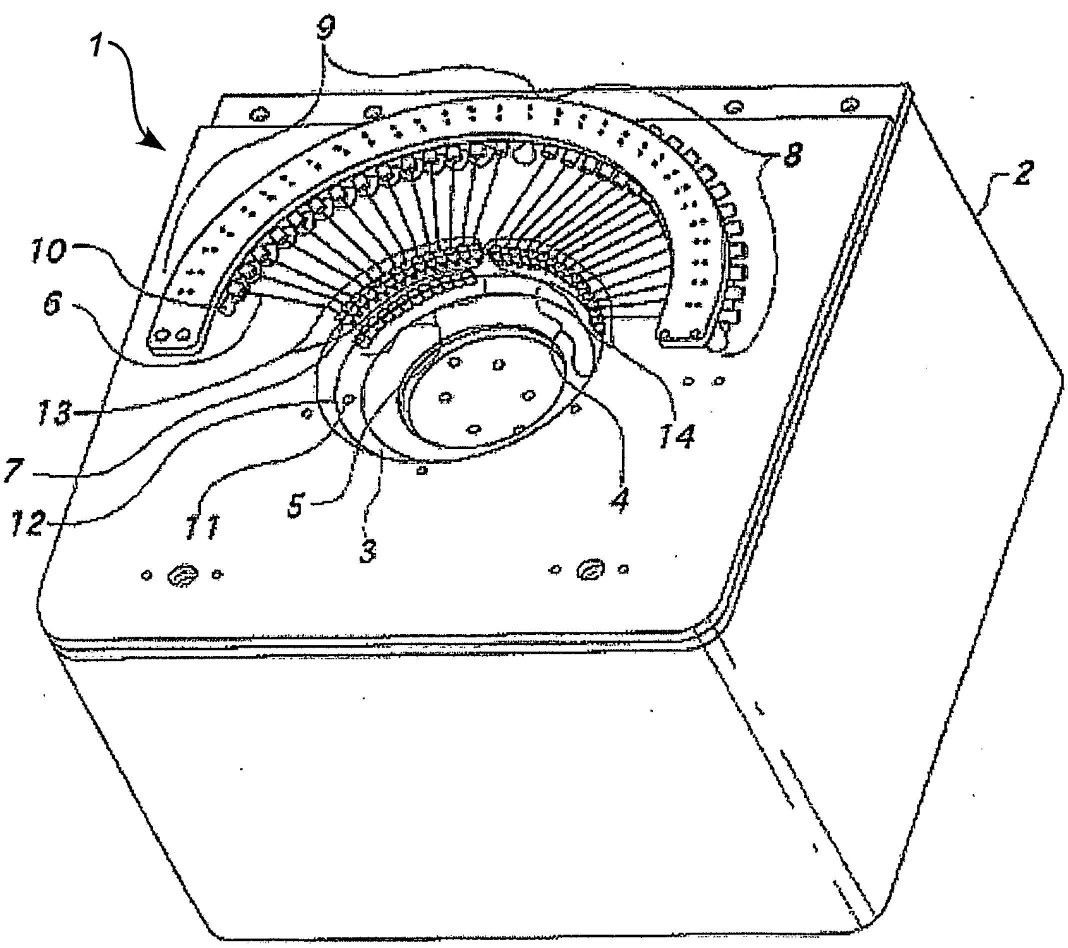


FIGURE 33

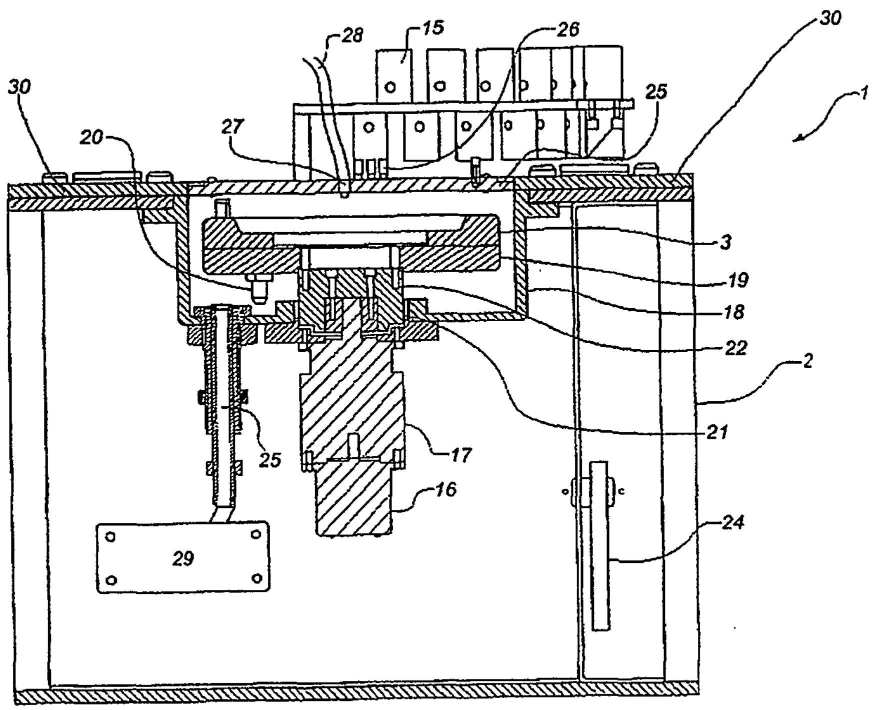


FIGURE 34

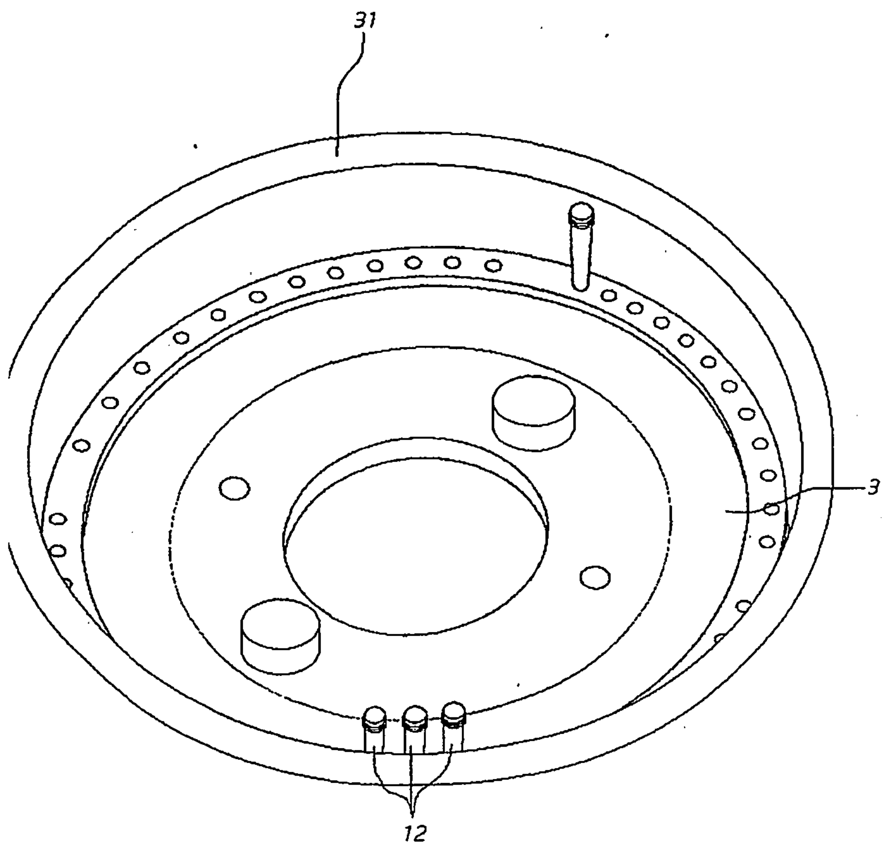


FIGURE 35

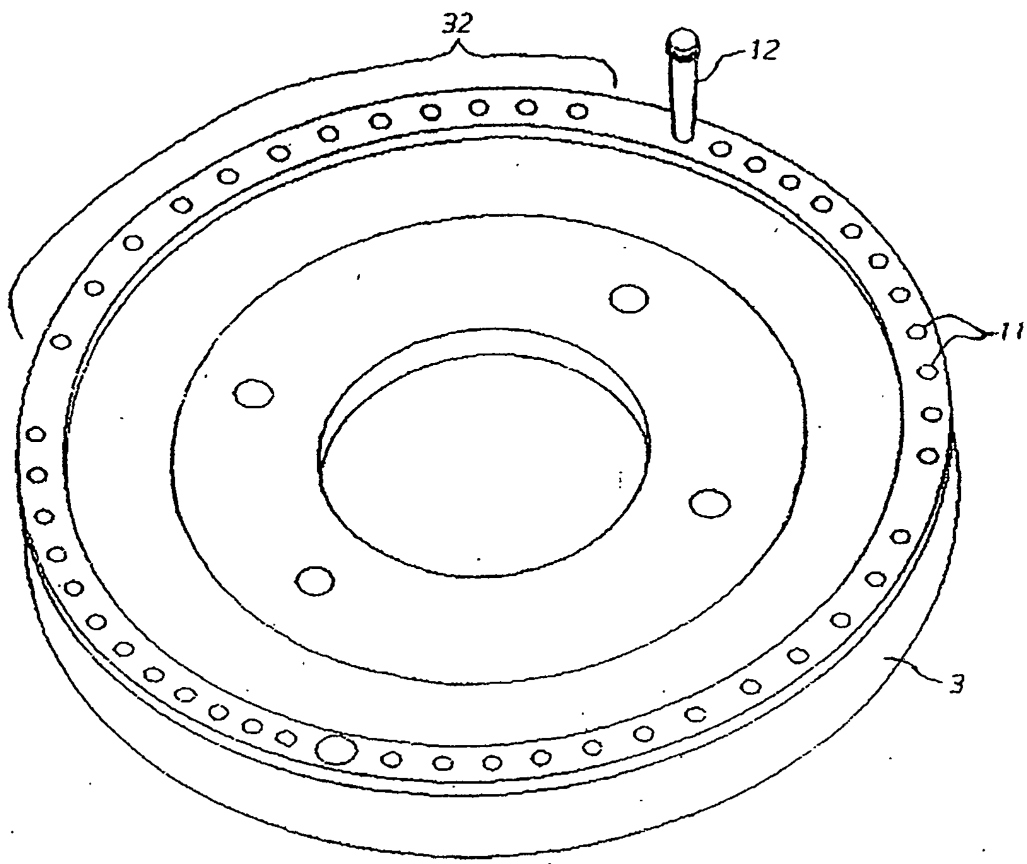


FIGURE 36

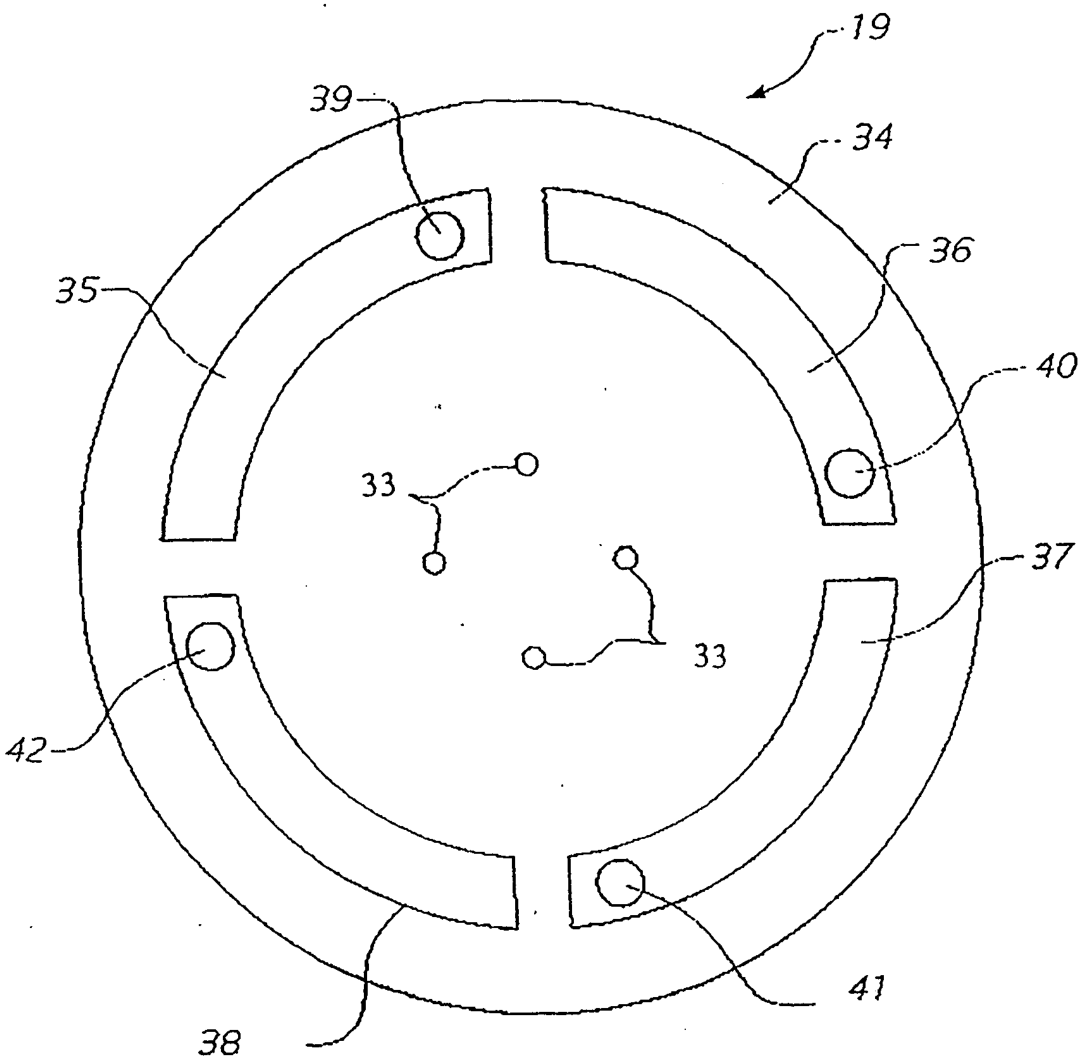
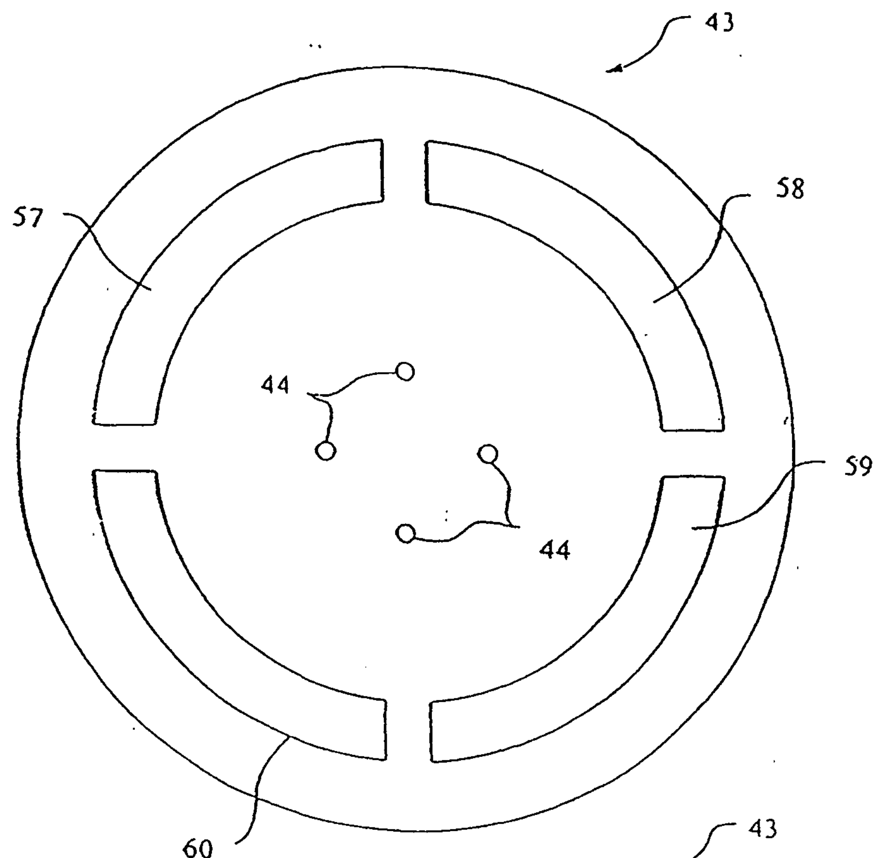


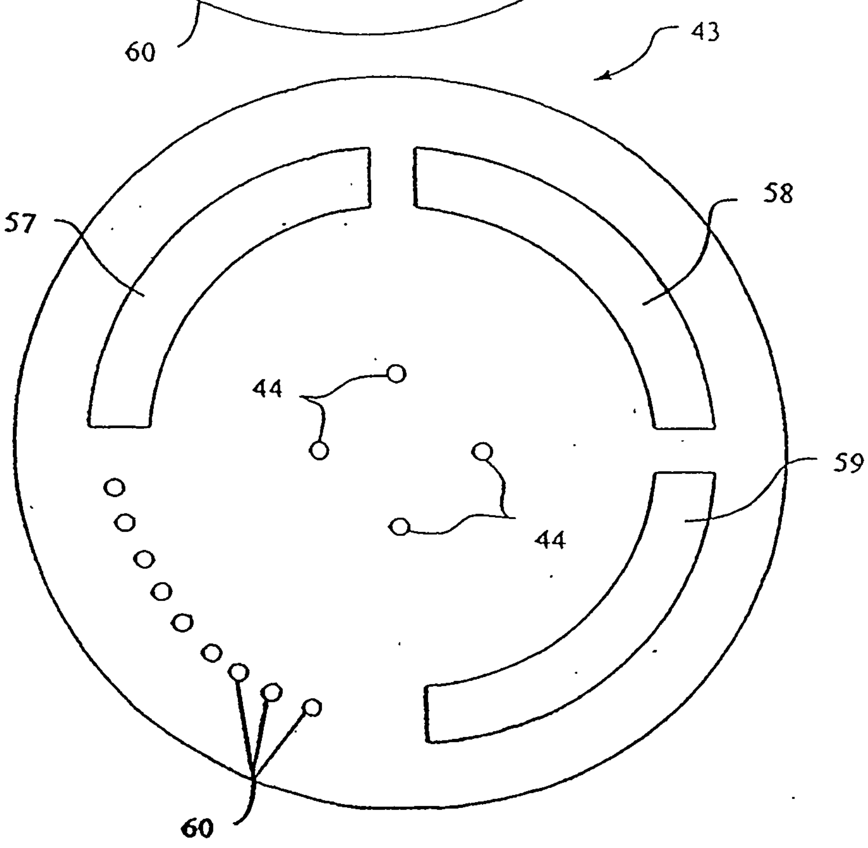
FIGURE 37

FIGURE 38

A



B



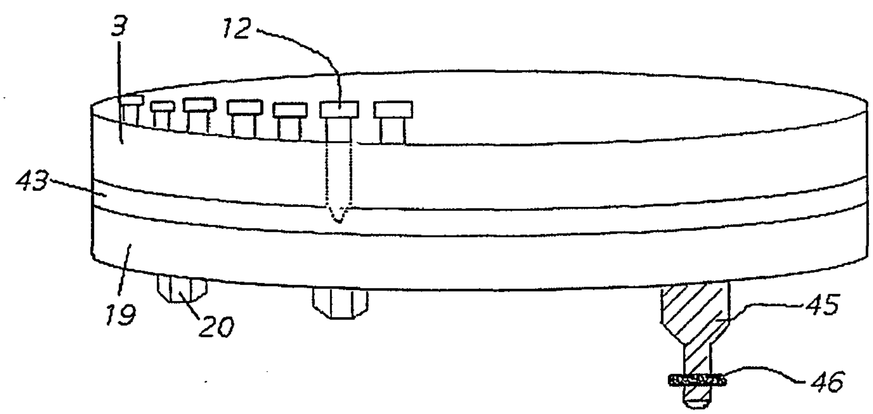


FIGURE 39

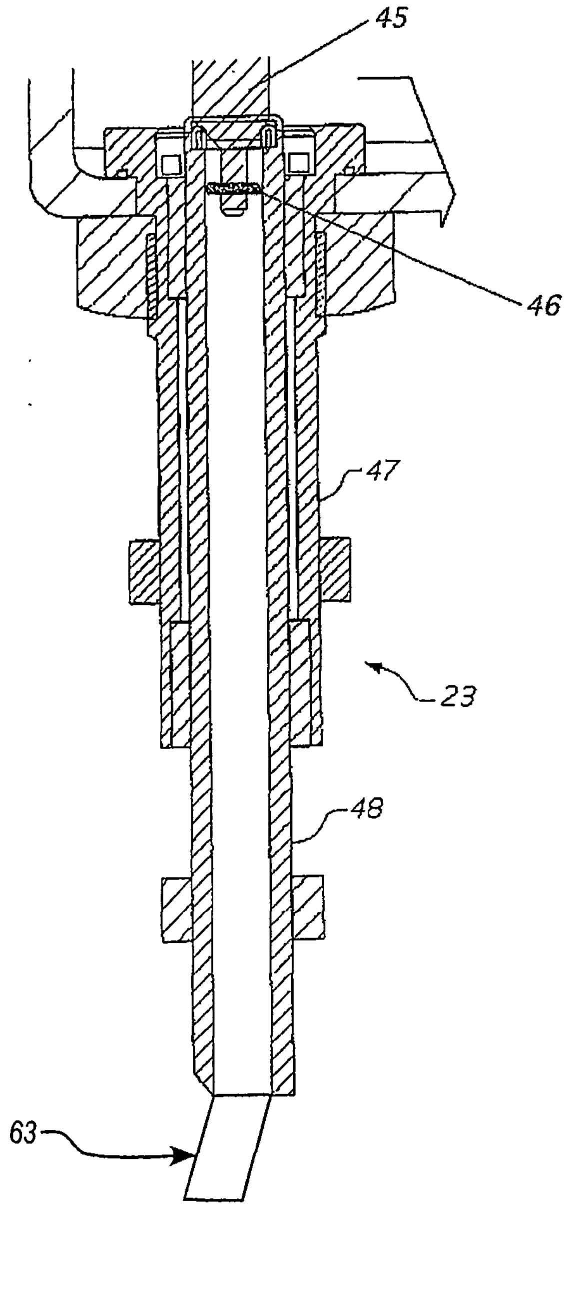
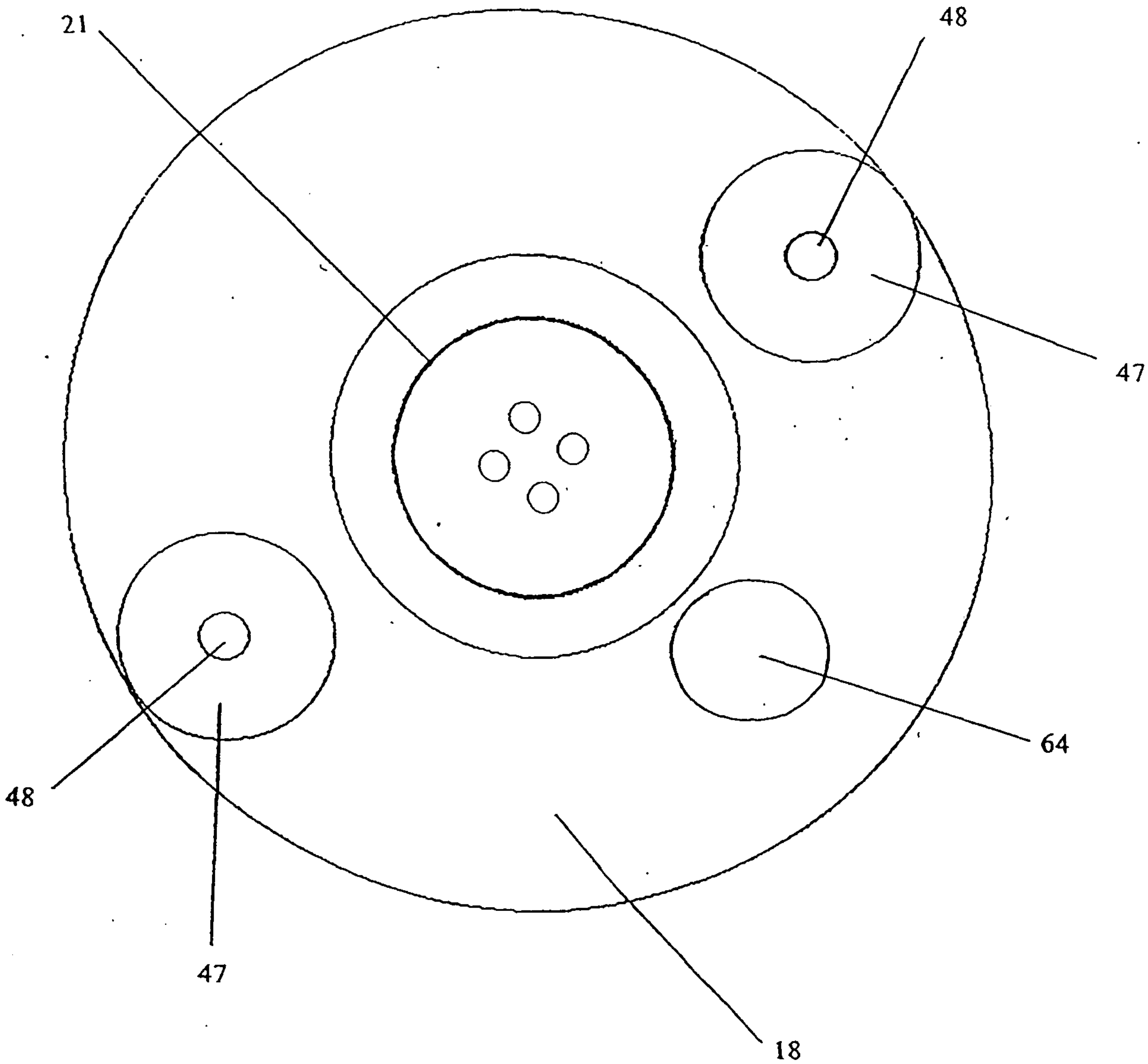


FIGURE 40

FIGURE 41



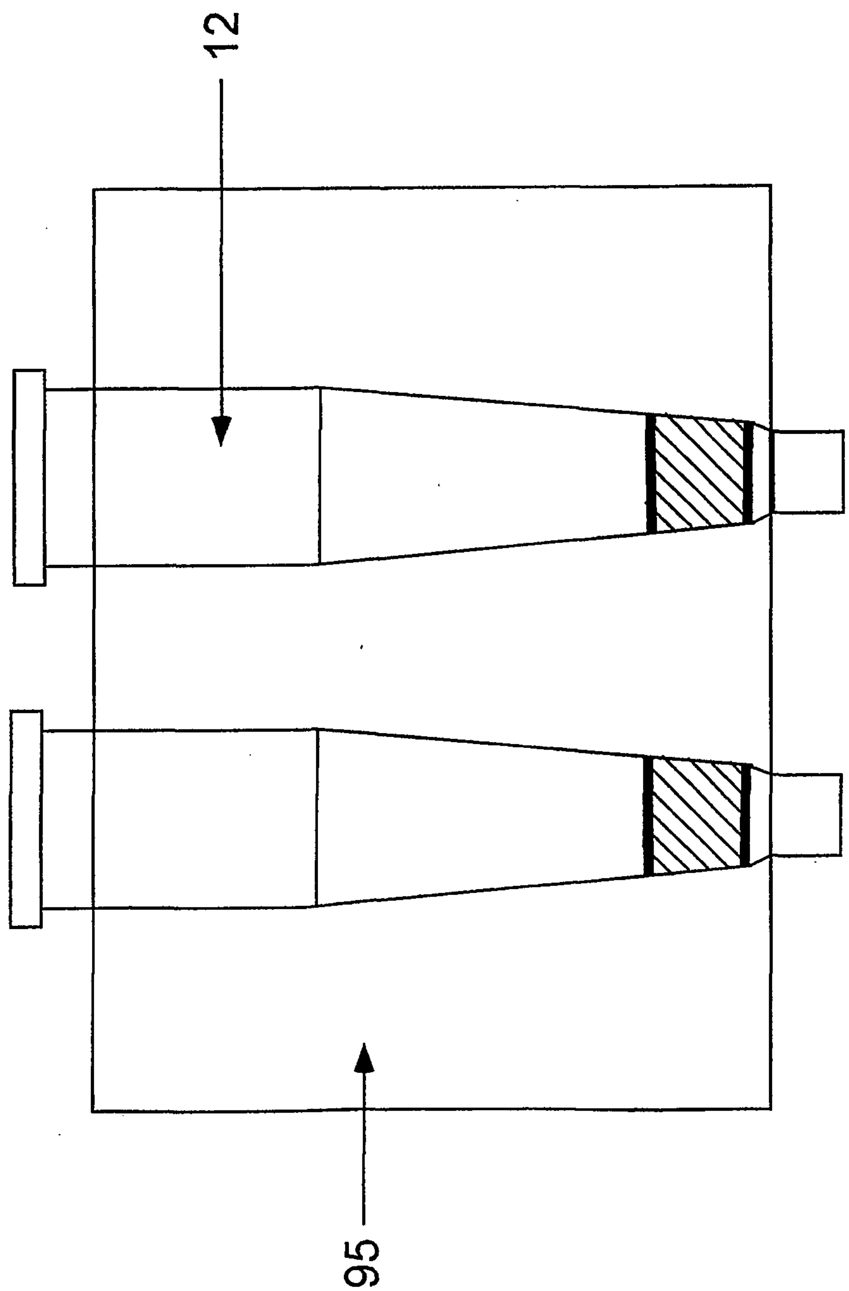


FIGURE 42A

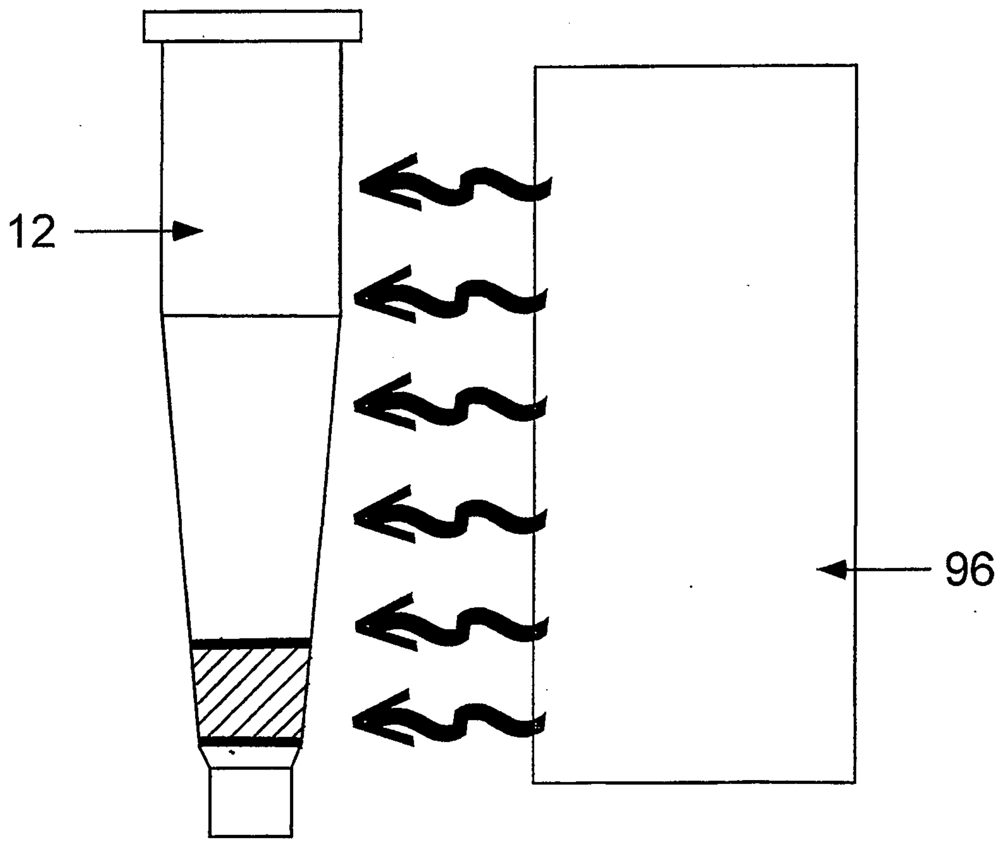


FIGURE 42B

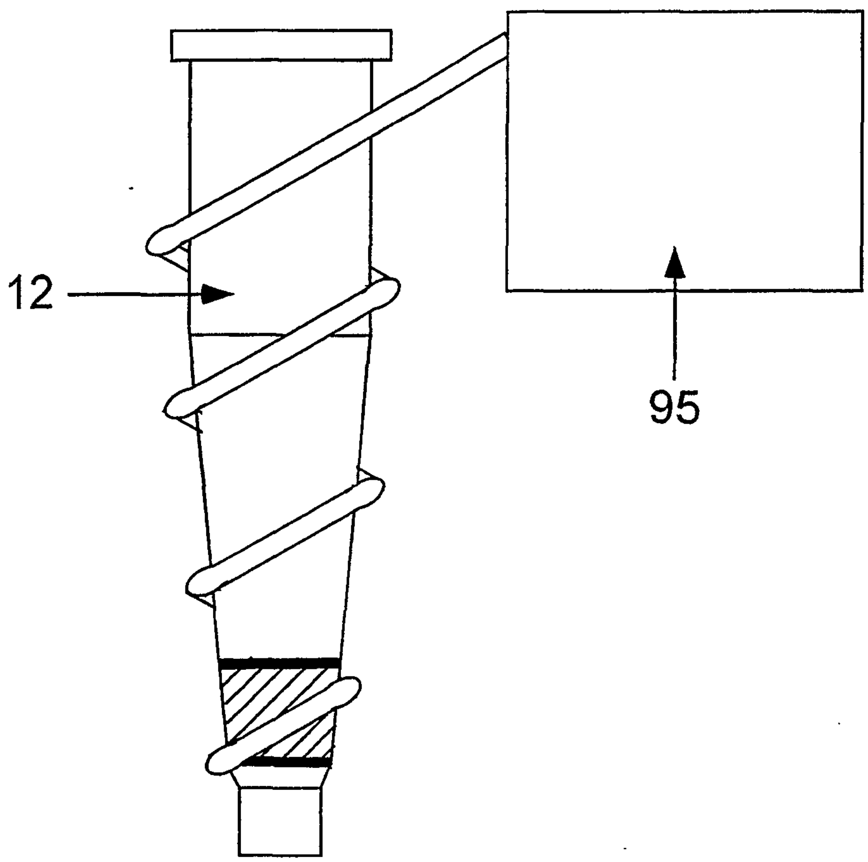


FIGURE 42C

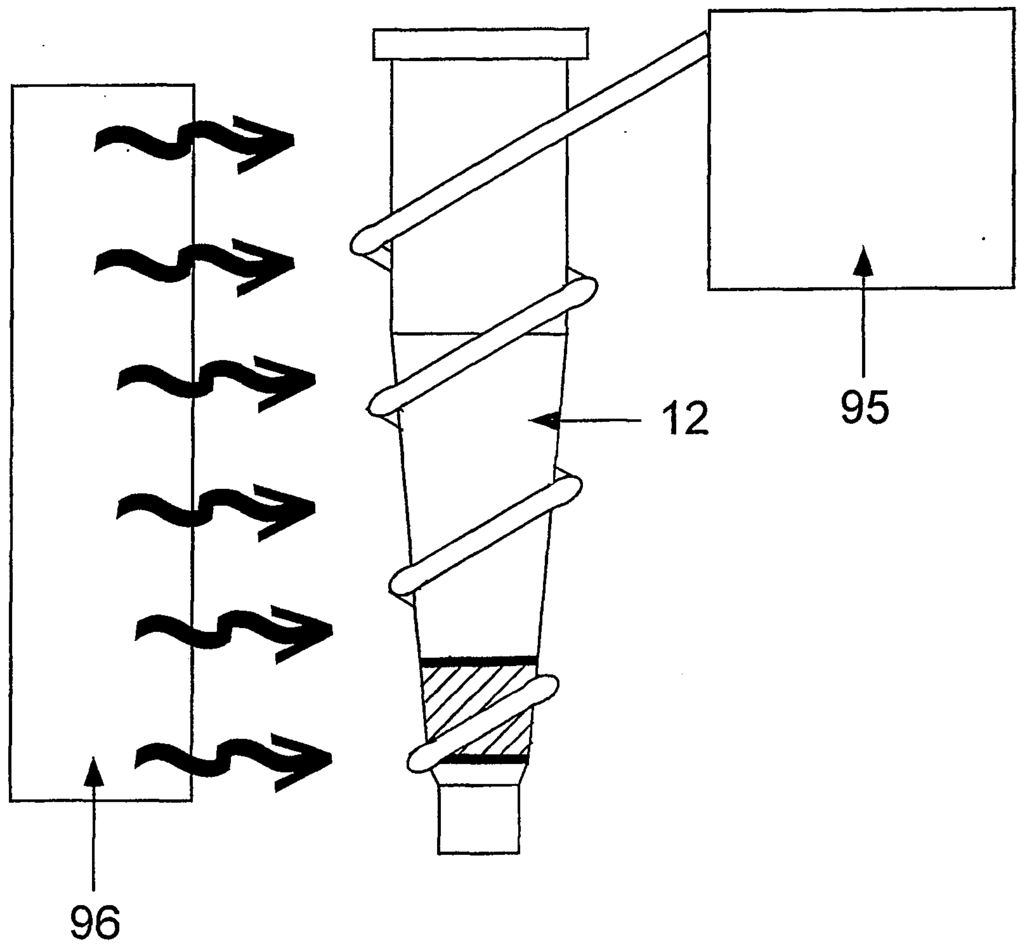


FIGURE 43A

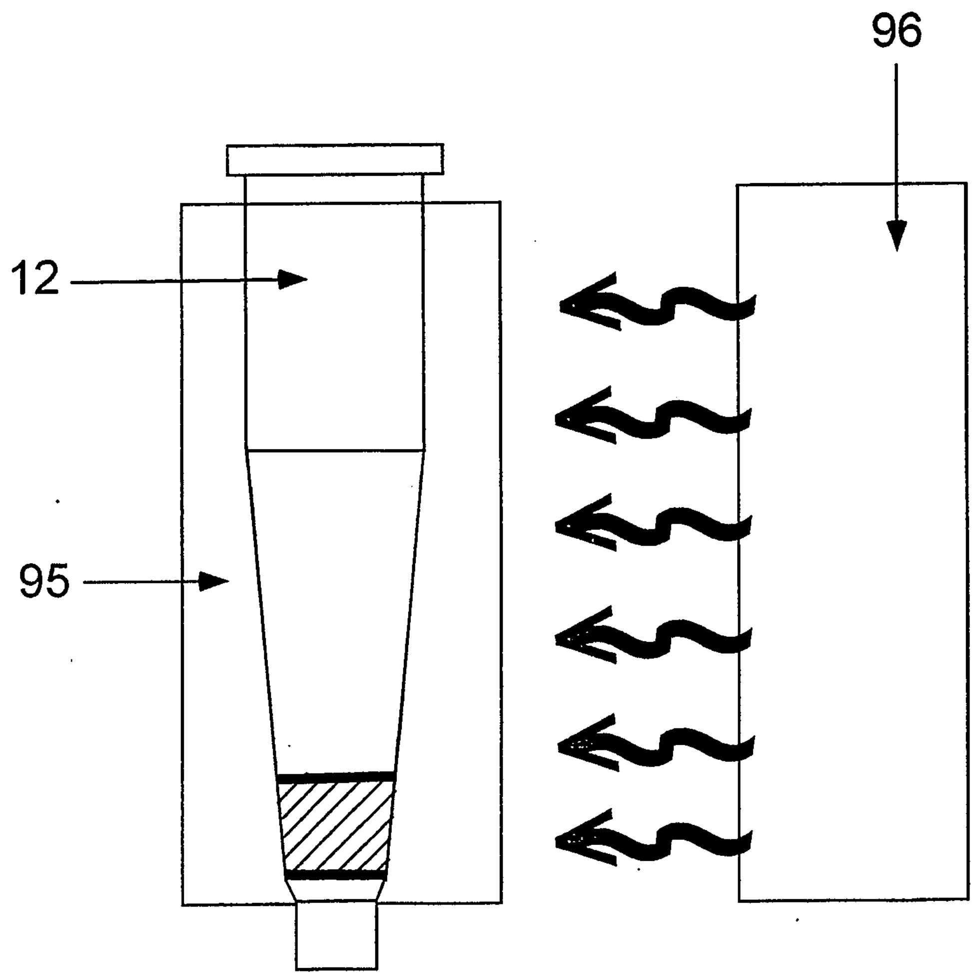


FIGURE 43B

FIGURE 44

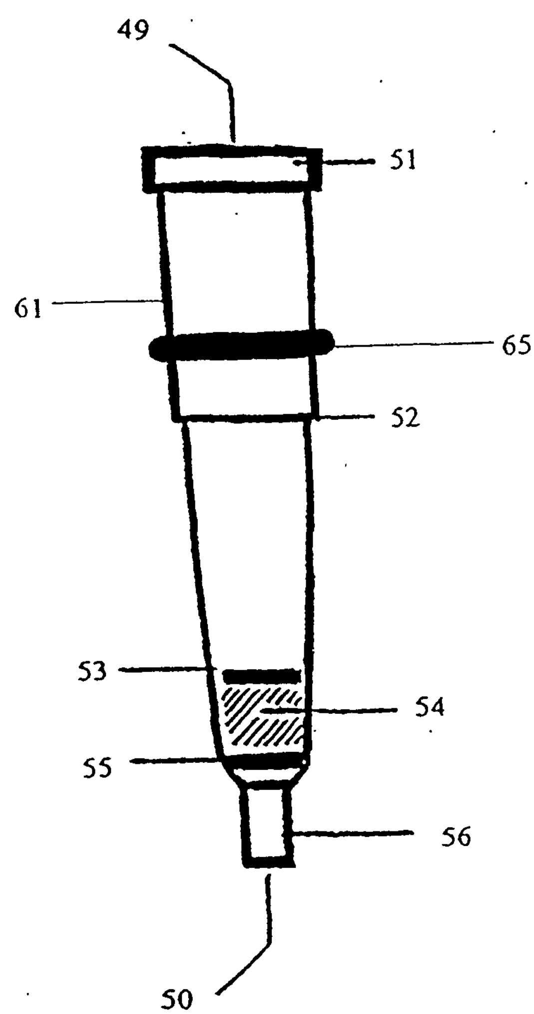


FIGURE 45

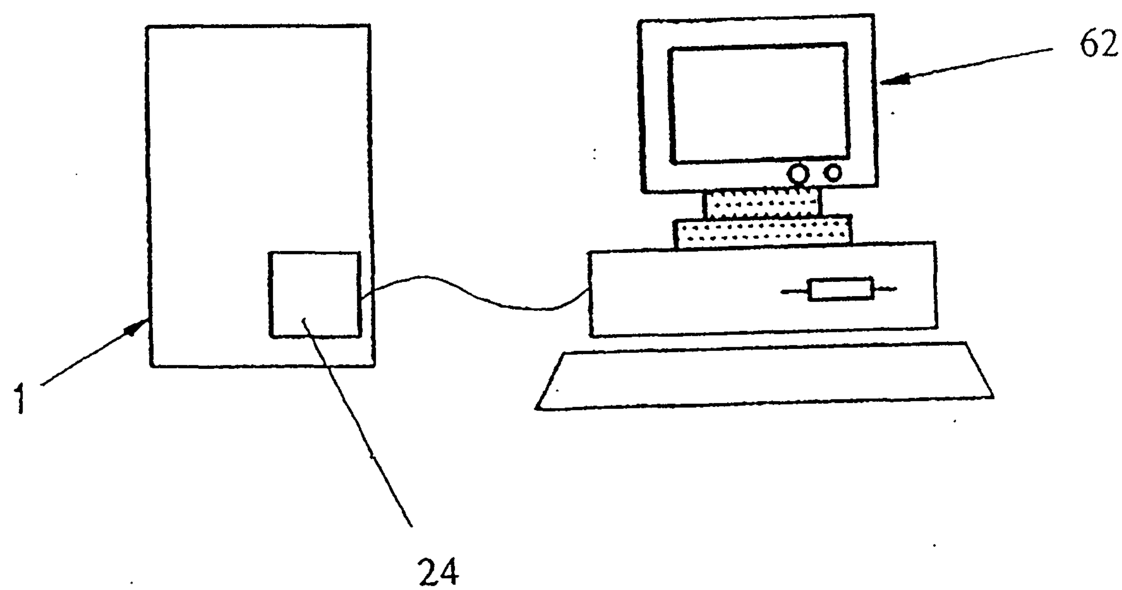


FIGURE 46A

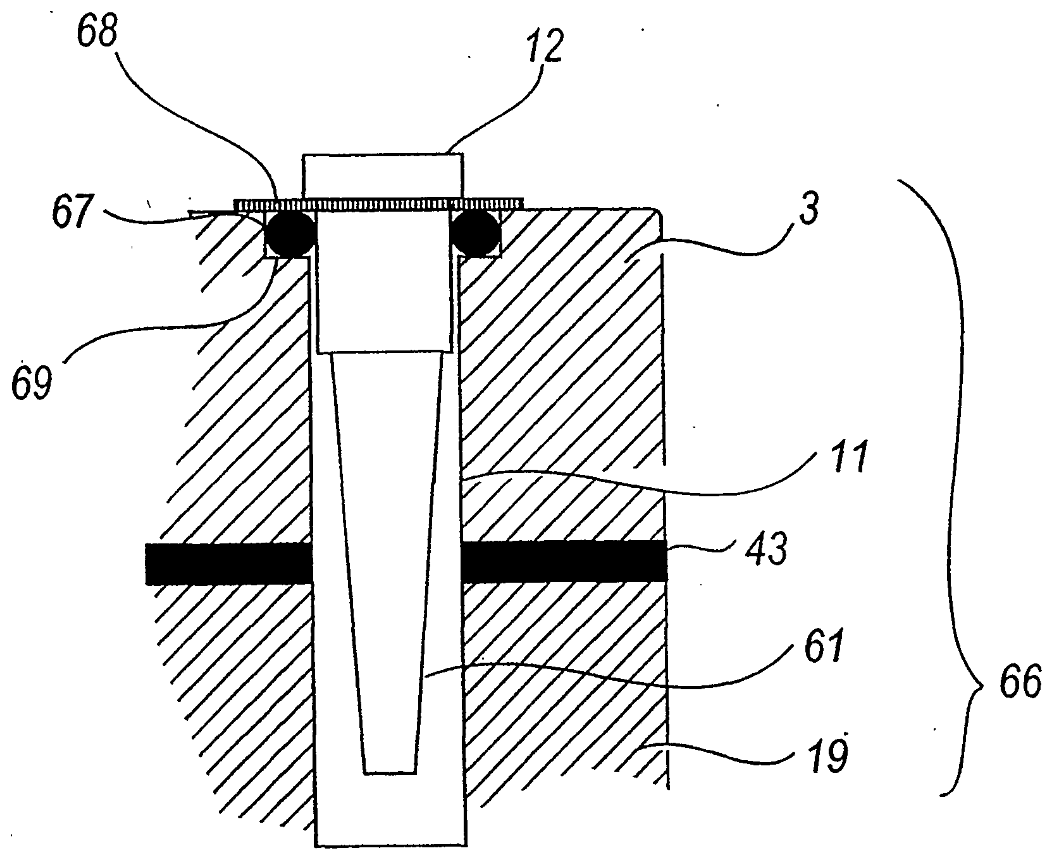


FIGURE 46B

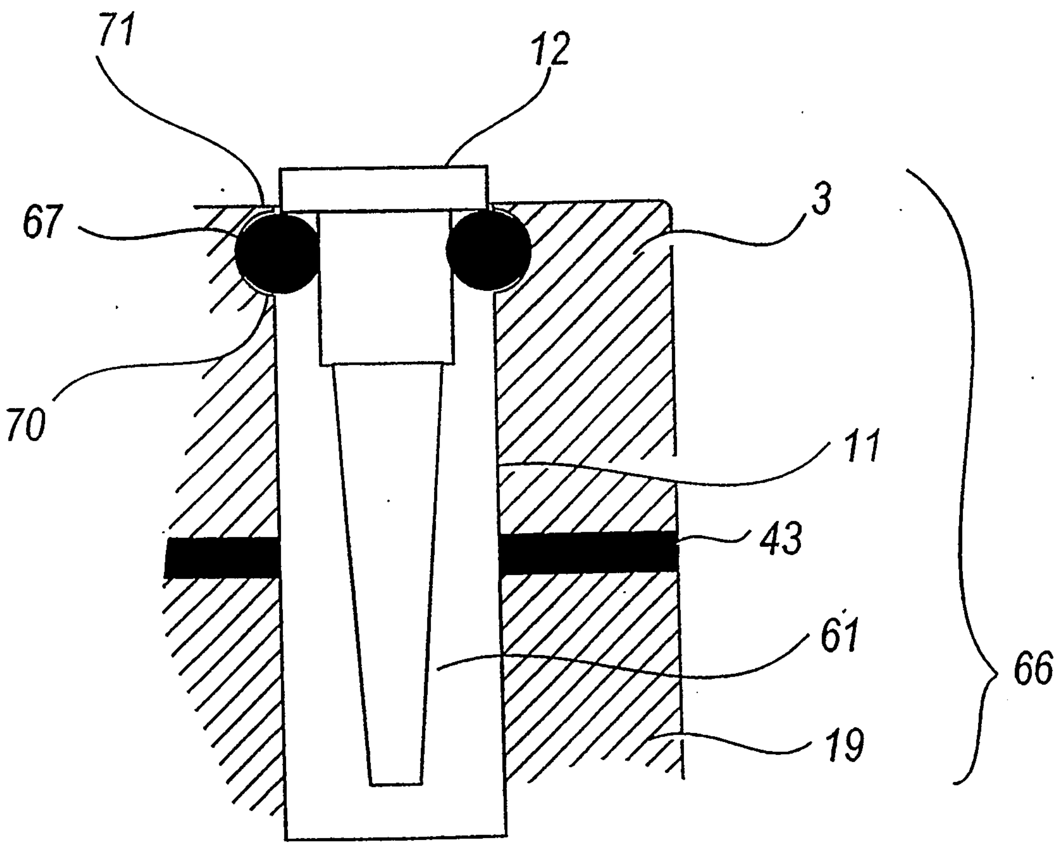


FIGURE 46C

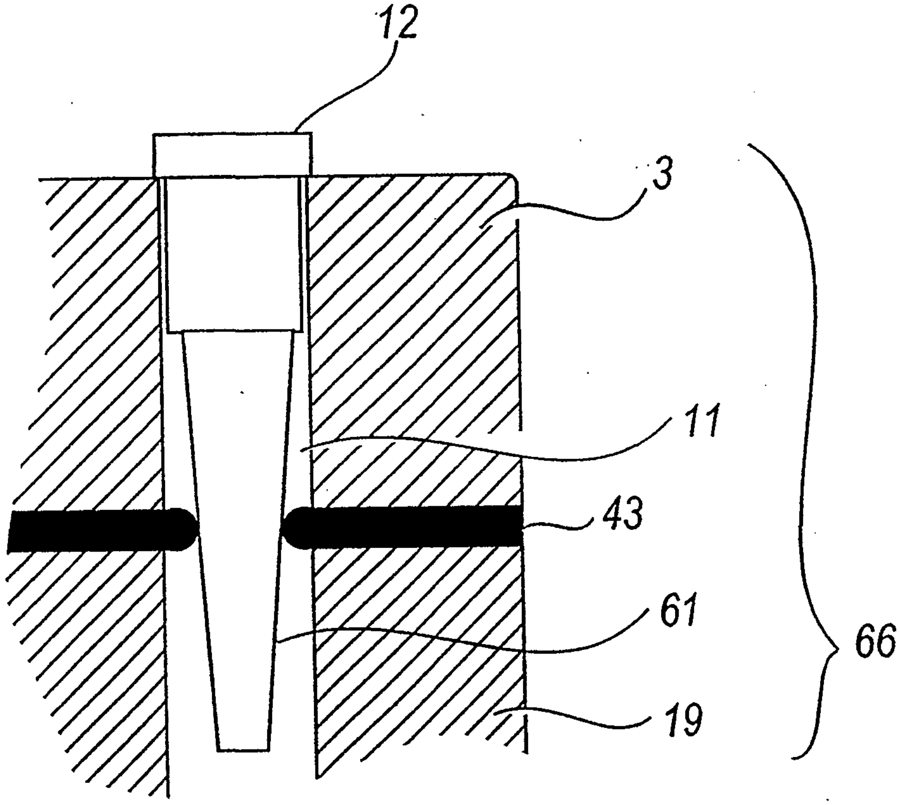


FIGURE 47A

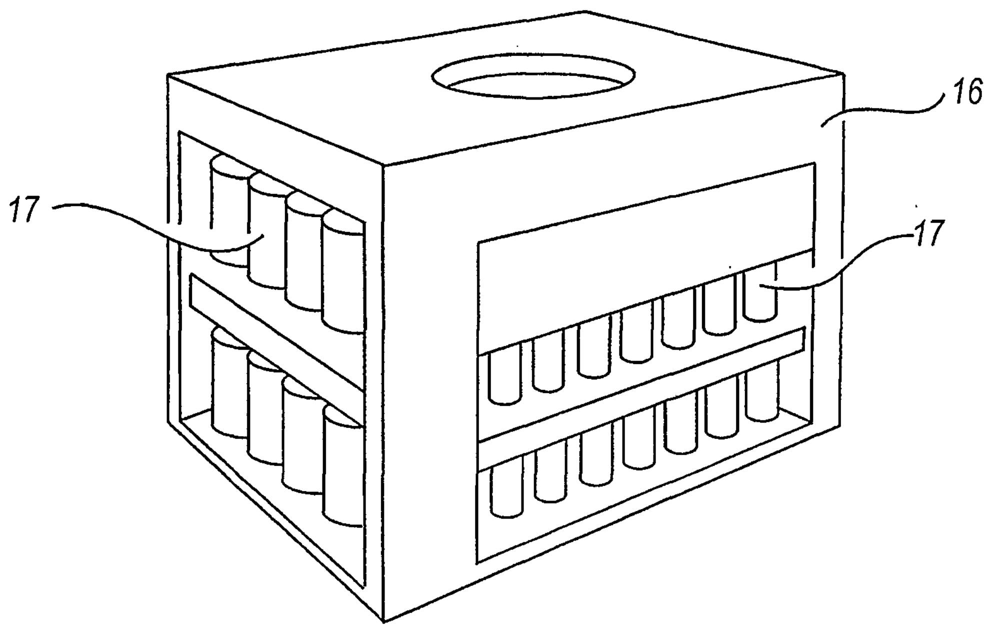


FIGURE 47B

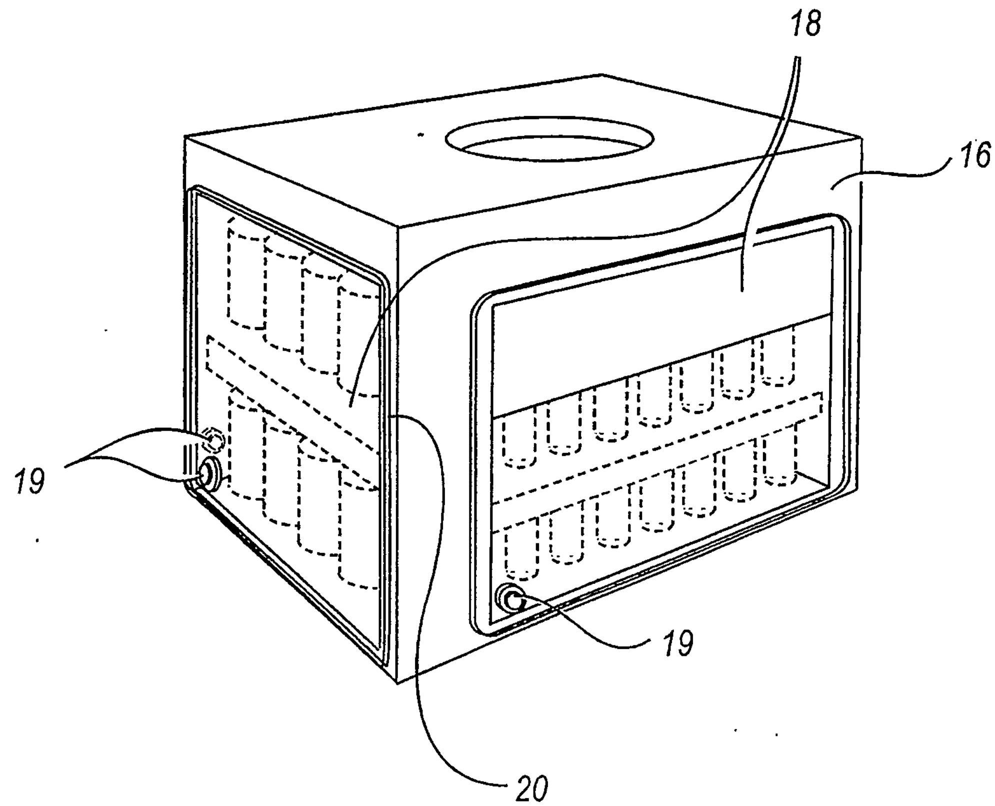


FIGURE 48A

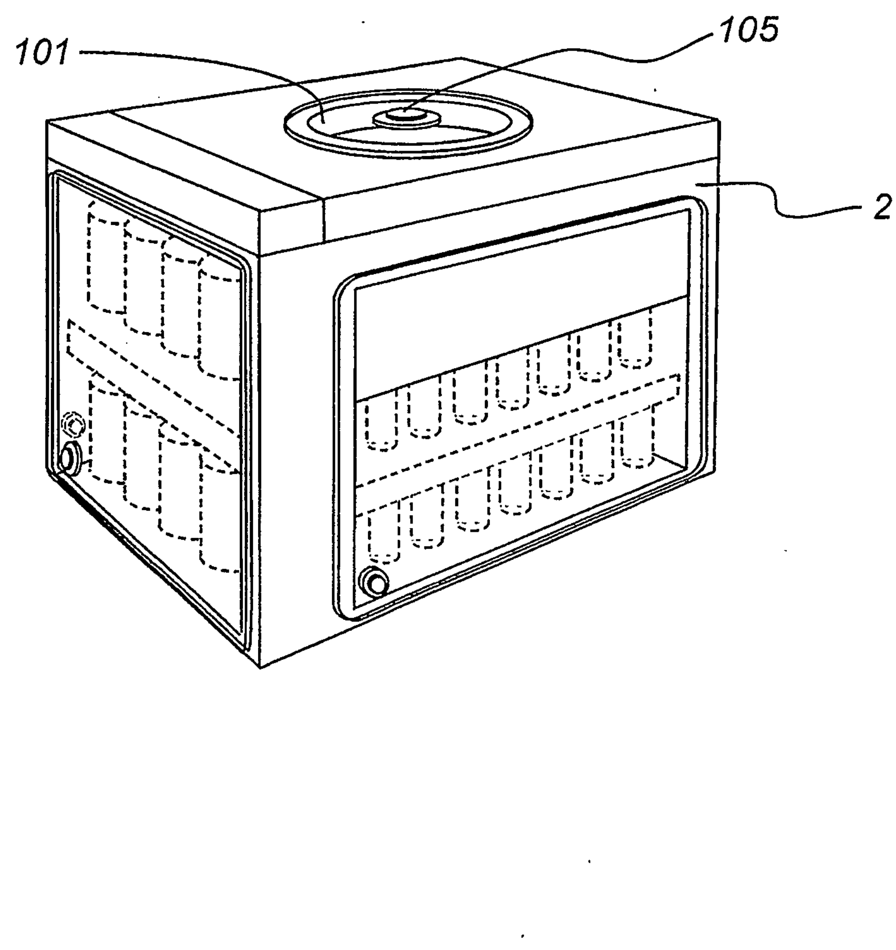


FIGURE 48B

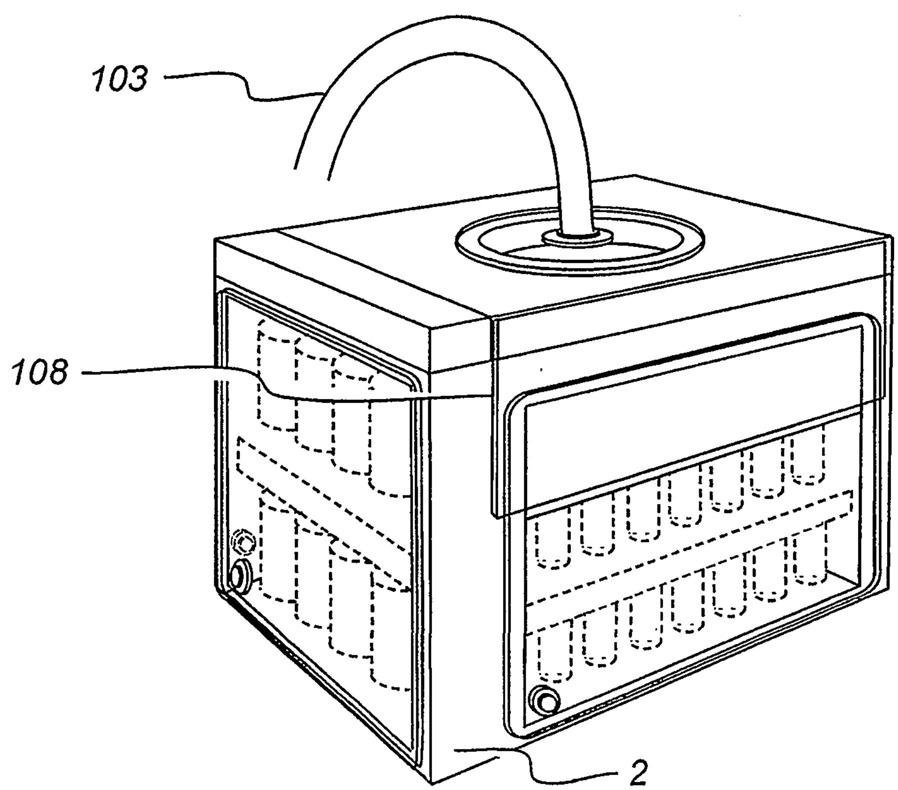


FIGURE 48C

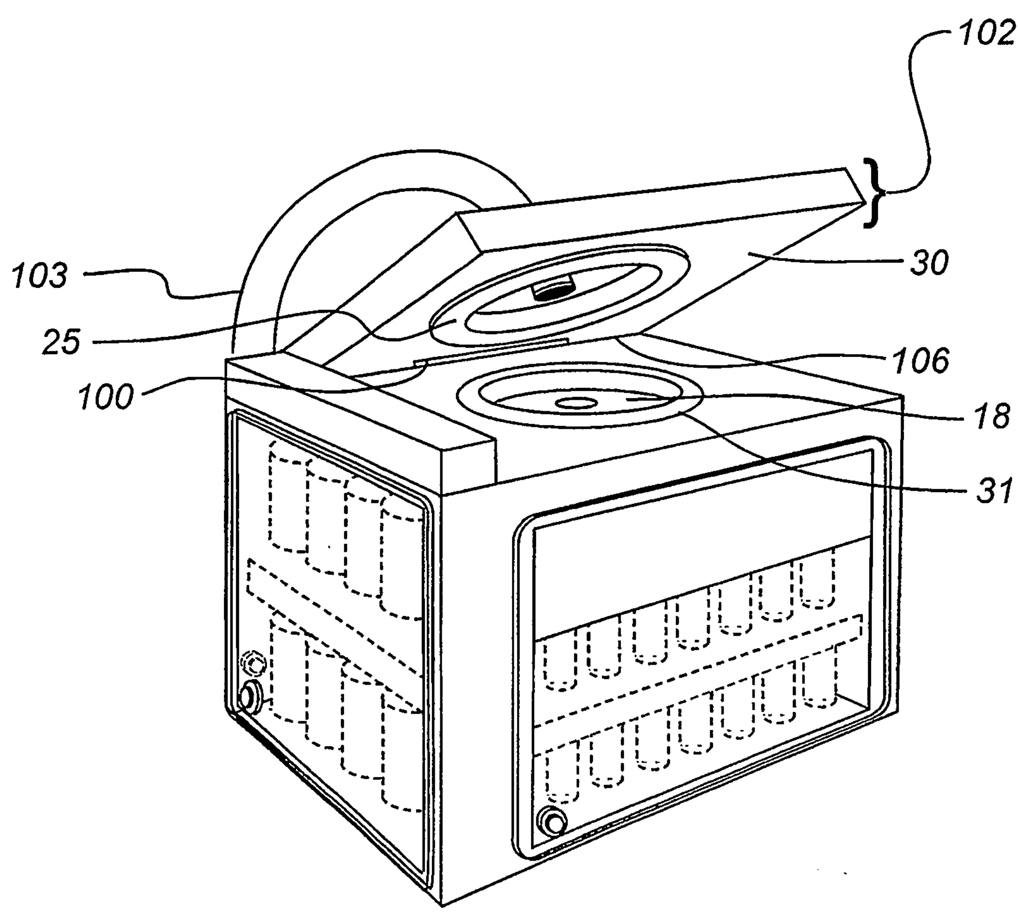


FIGURE 49A

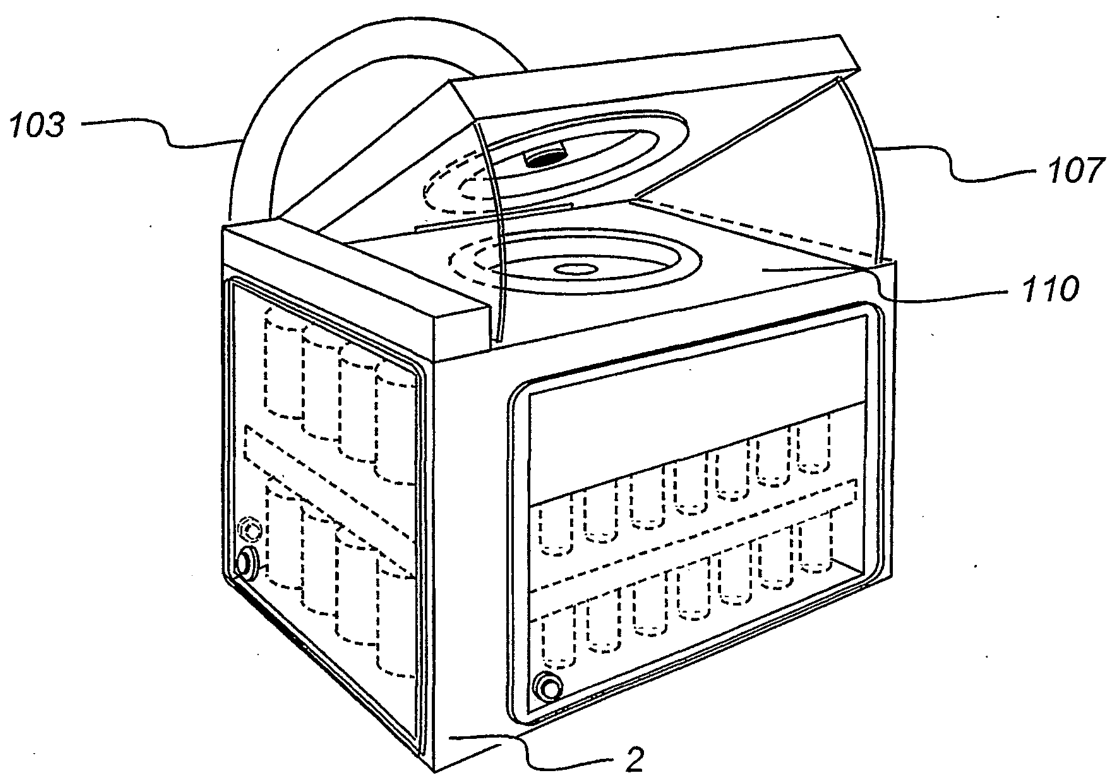


FIGURE 49B

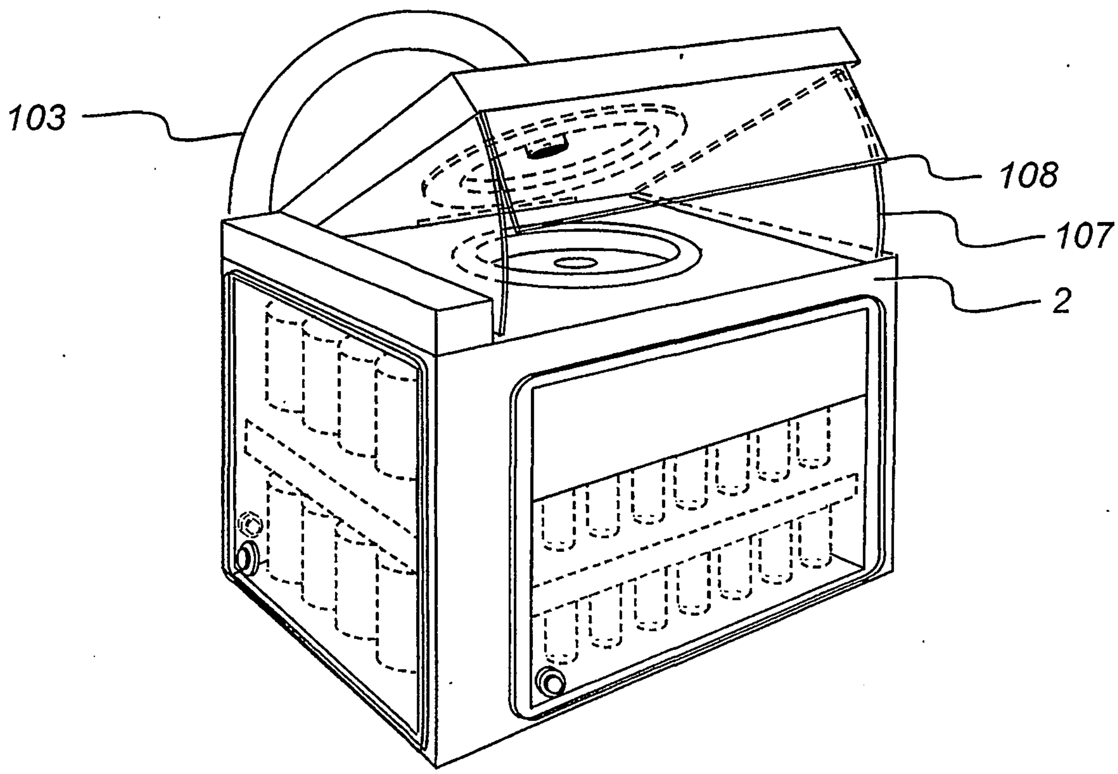


FIGURE 49C

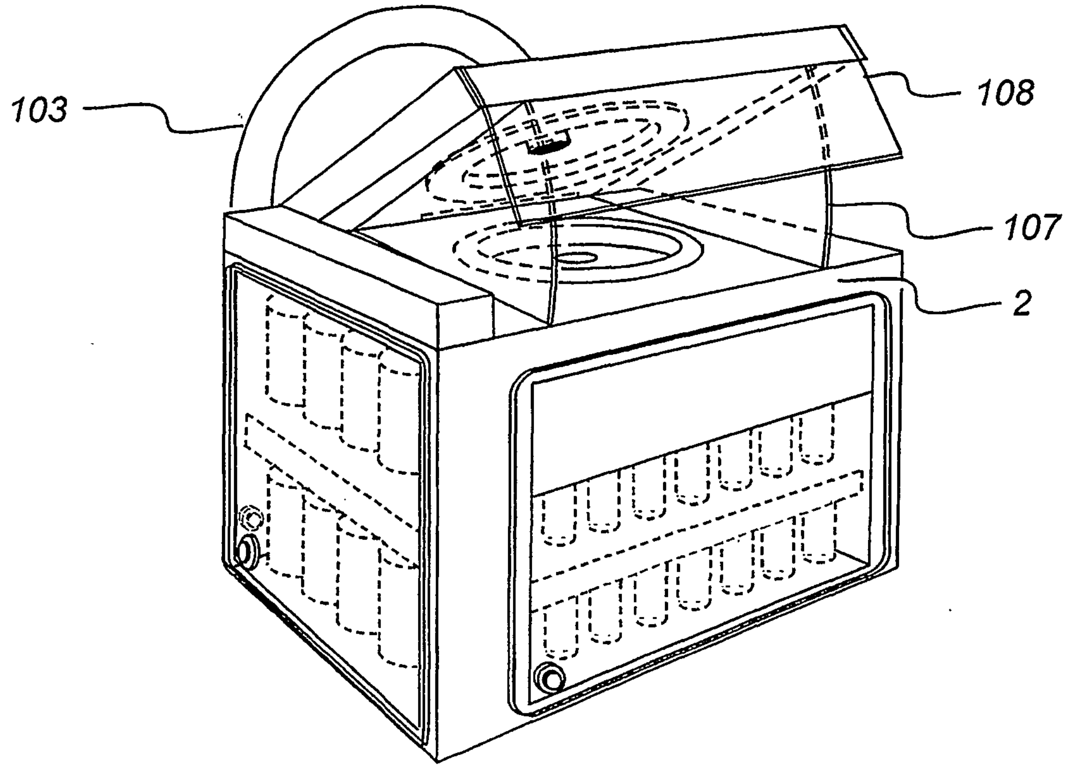


FIGURE 50

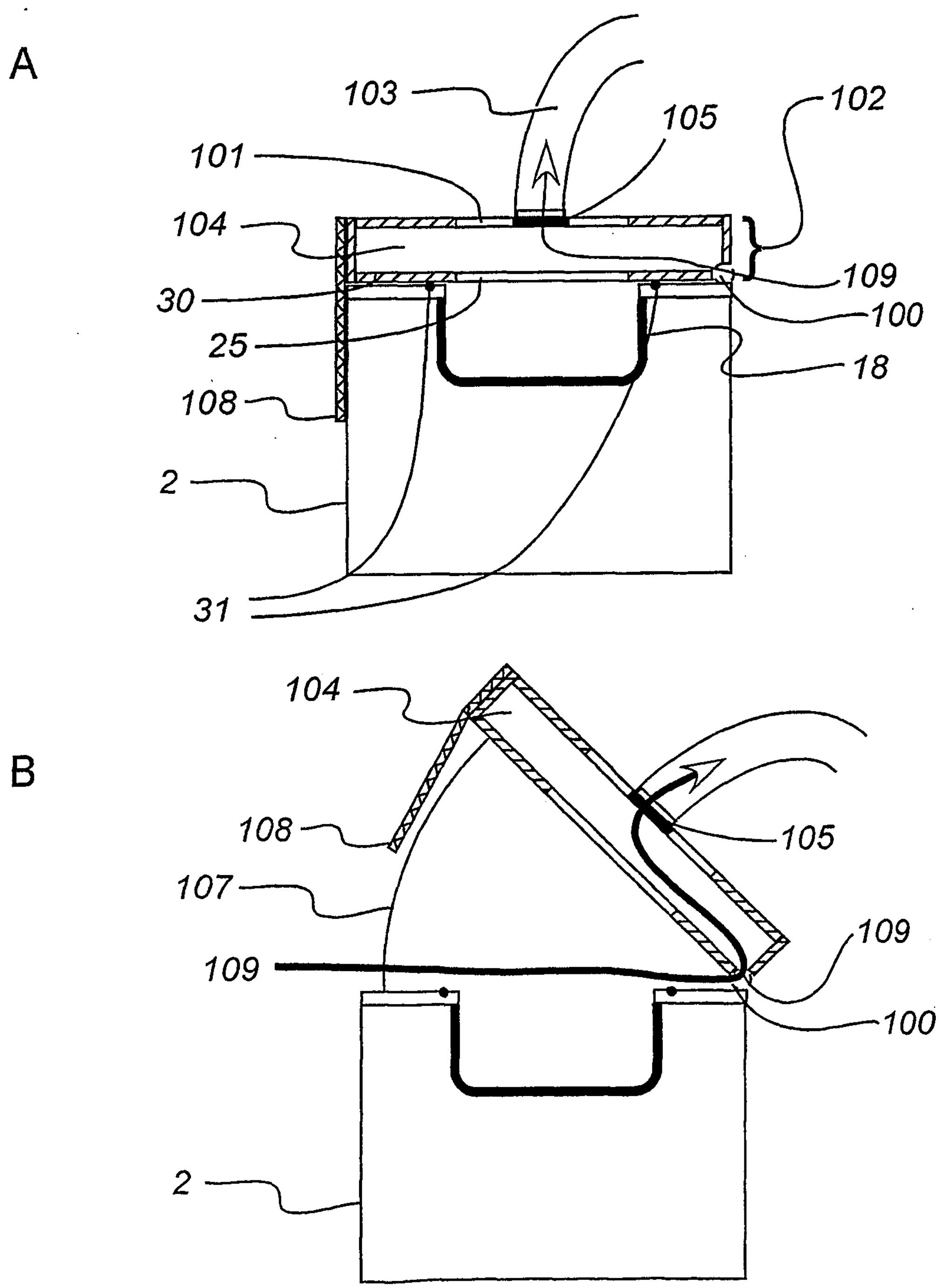
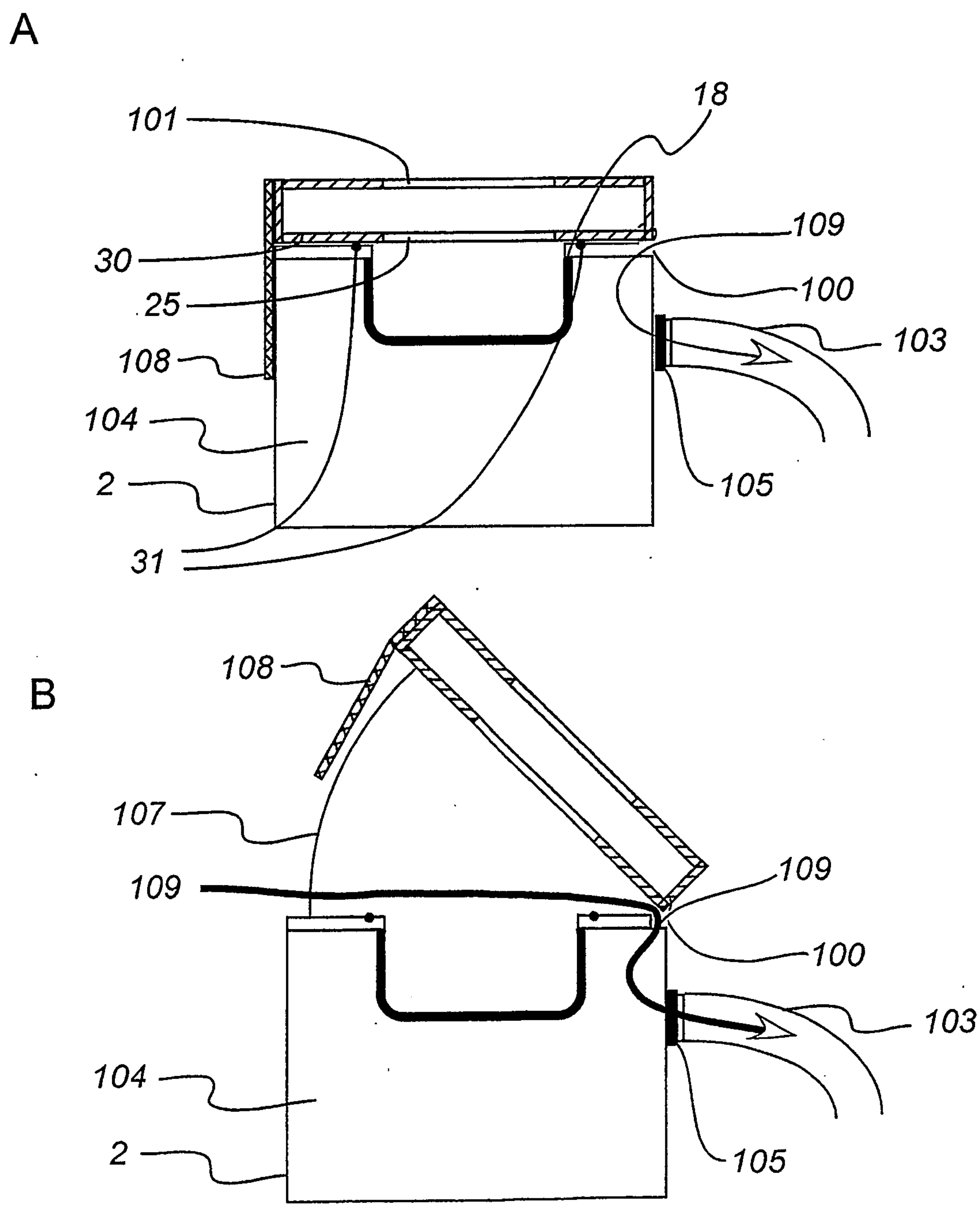

A
103
105
102
101
104
109
30
100
25
18
108
2
31
B
104
108
105
107
109
109
100
2

FIGURE 51


A
101
18
109
30
25
100
103
108
104
2
105
31
B
108
107
109
109
100
103
104
2
105

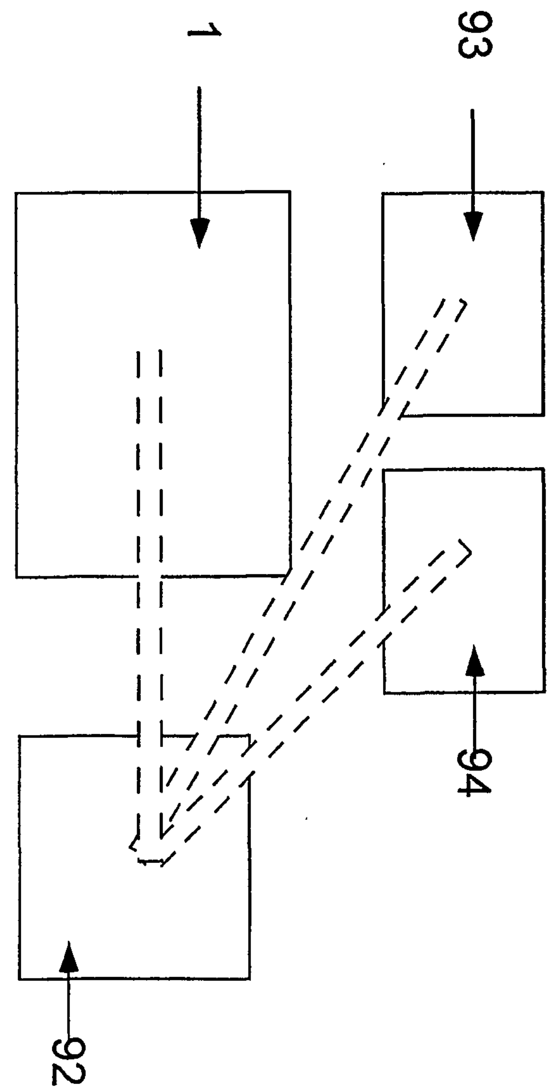


FIGURE 52

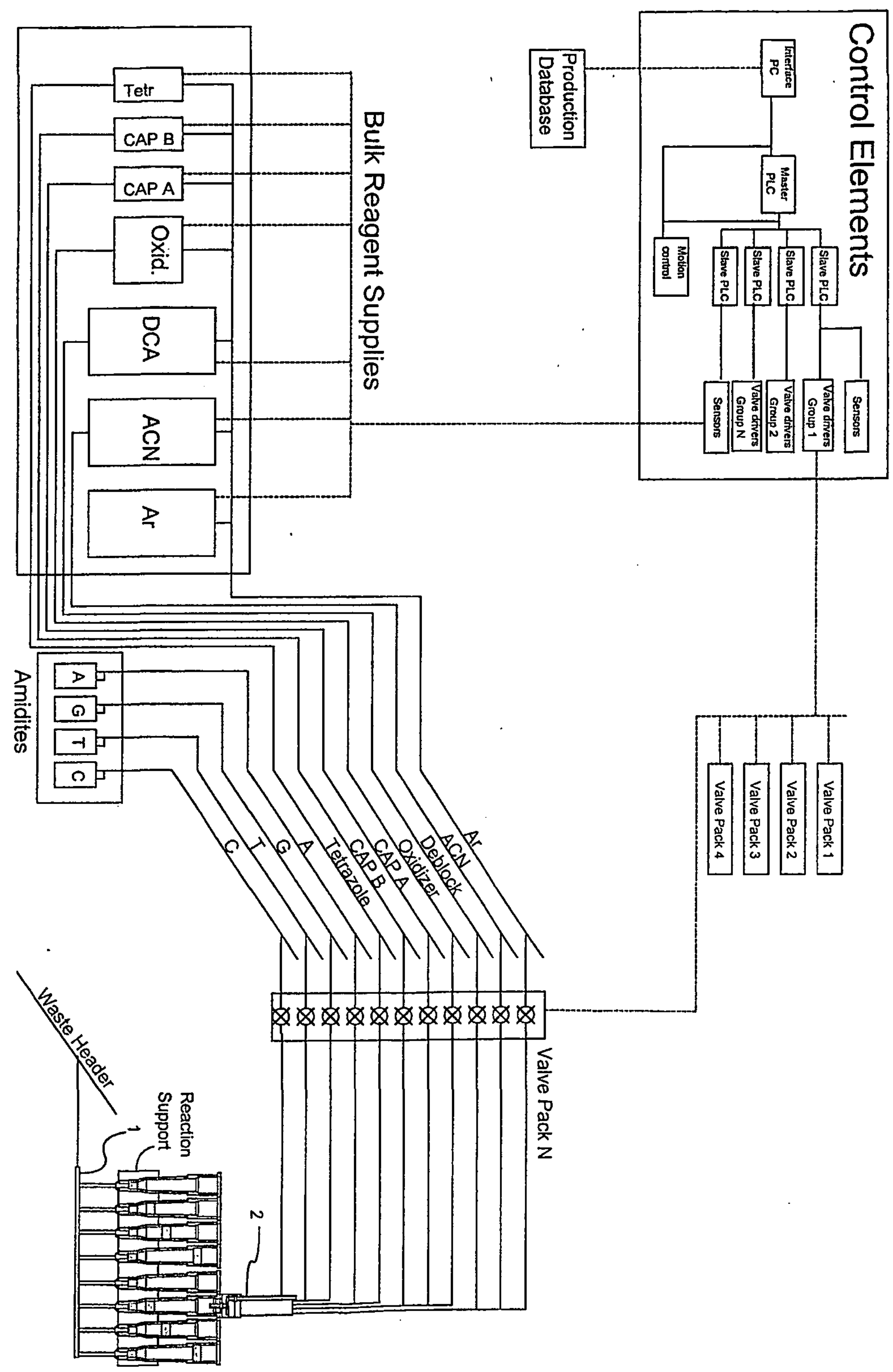

Control Elements
Interface PC
Master PLC
Motion control
Slave PLC
Slave PLC
Slave PLC
Slave PLC
Sensors
Valve drivers Group 1
Valve drivers Group 2
Valve drivers Group N
Sensors
Production Database
Bulk Reagent Supplies
Tetr
CAP B
CAP A
Oxid.
DCA
ACN
Ar
Amidites
A
G
T
C
Valve Pack 1
Valve Pack 2
Valve Pack 3
Valve Pack 4
Ar
ACN
Deblock
Oxidizer
CAP A
CAP B
Tetrazole
A
G
T
C
Valve Pack N
Waste Header
Reaction Support
1
2

FIGURE 53

A

B

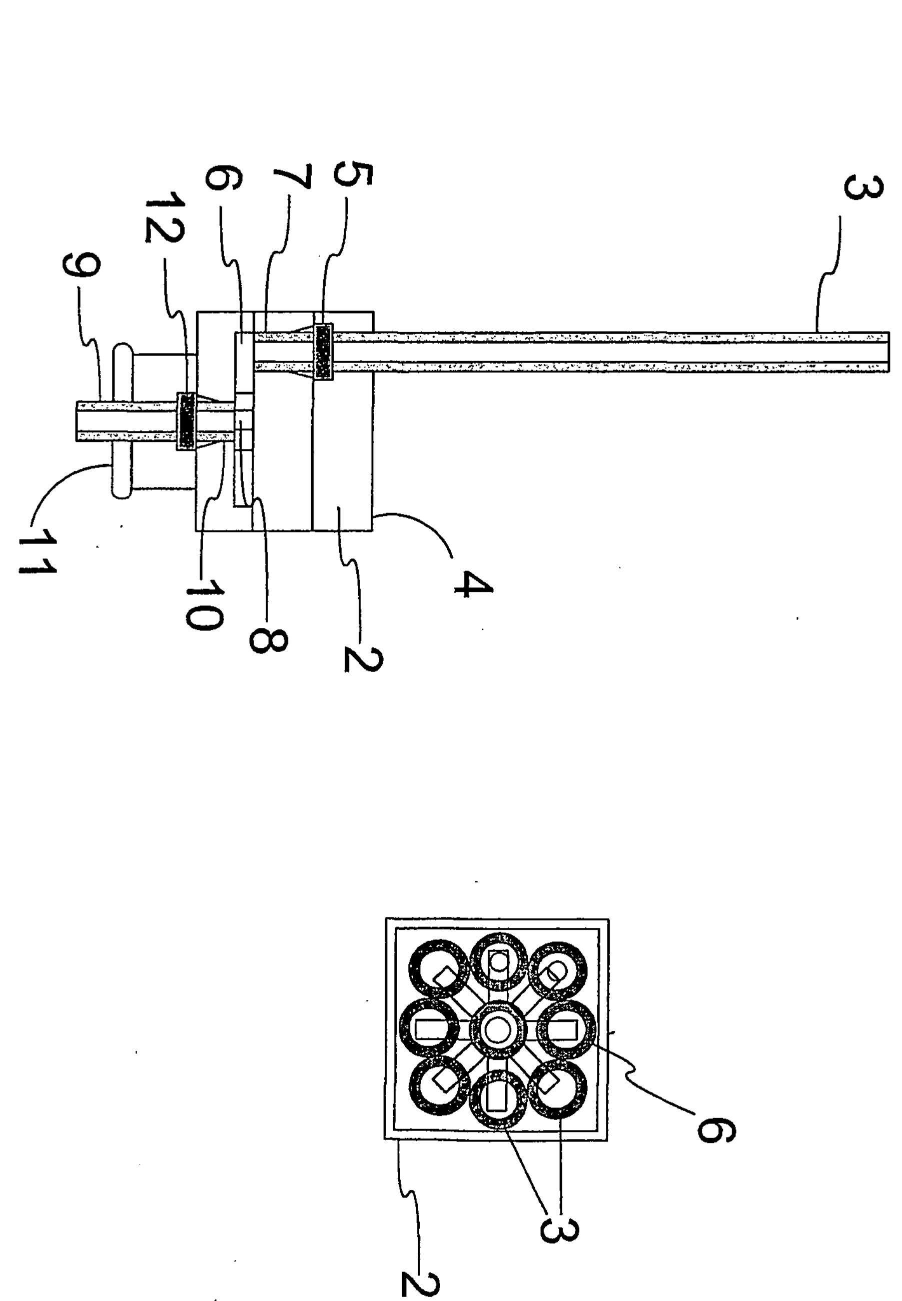


FIGURE 54

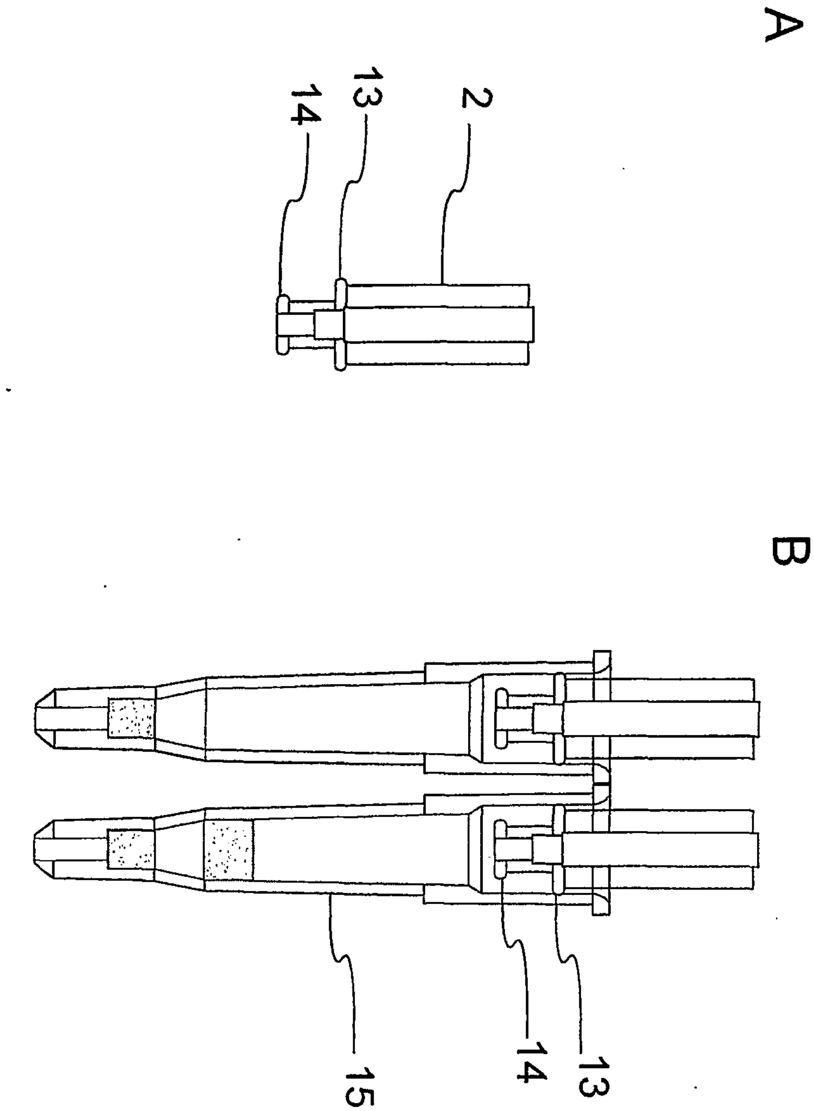


FIGURE 55

FIGURE 56

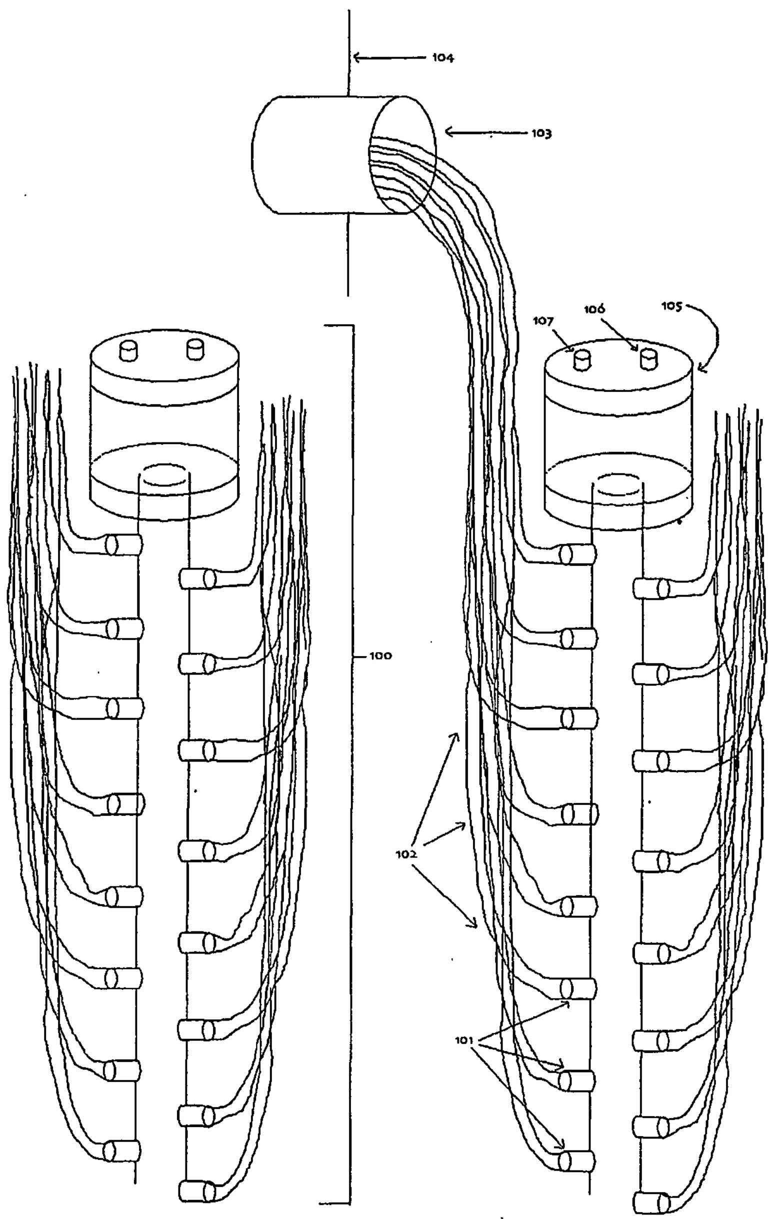


FIGURE 57

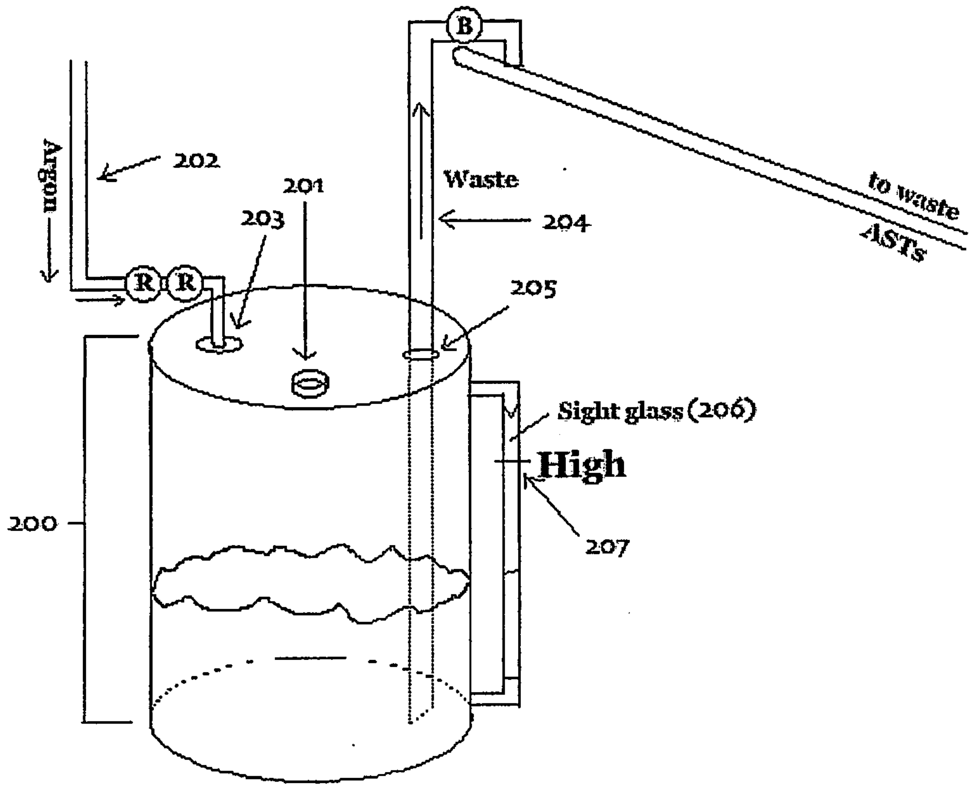

B
202
Argon
R R
201
203
Waste
204
to waste
ASTs
205
Sight glass (206)
High
200
207

Figure 58a

Integrated Data & Process Flows

Generic Overview - Modules and Interfaces

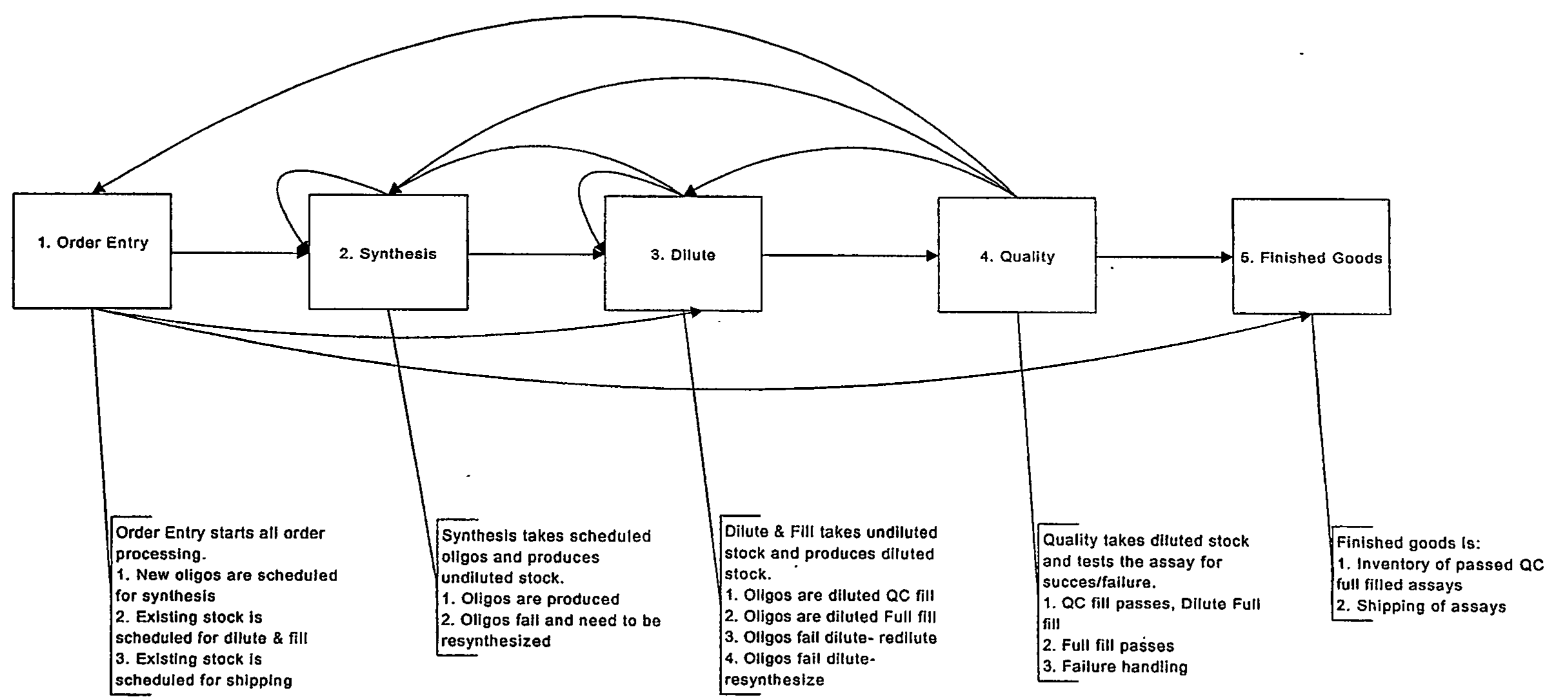


Figure 58b

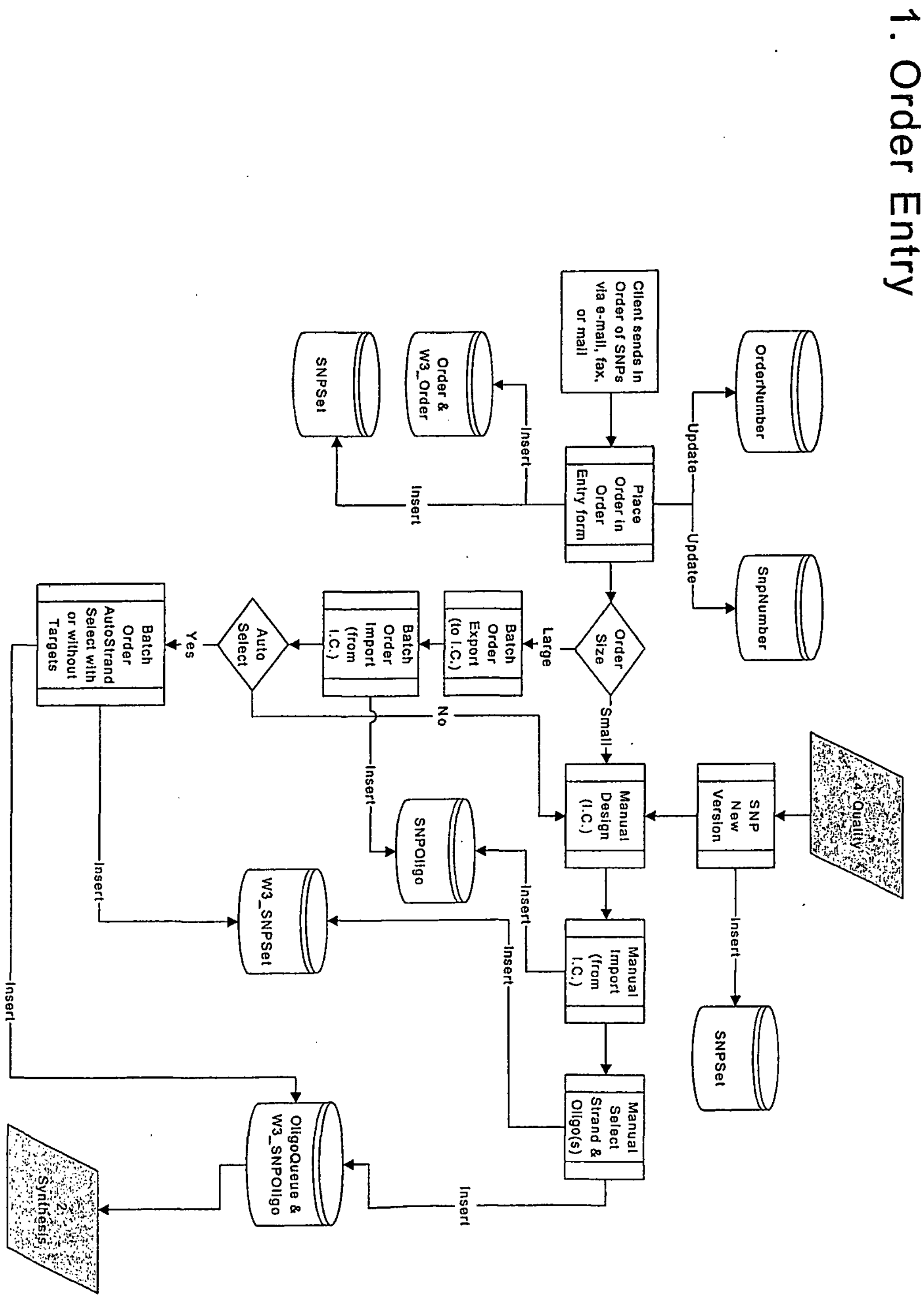

1. Order Entry
Client sends in Order of SNPs via e-mail, fax, or mail
Place Order in Order Entry form
OrderNumber
SnpNumber
Update
Update
Insert
Insert
Order & W3_Order
SNPSet
Order Size
Large
Small
Batch Order Export (to I.C.)
Batch Order Import (from I.C.)
Auto Select
Yes
No
Batch Order AutoStrand Select with or without Targets
4. Quality
SNP New Version
Insert
SNPSet
Manual Design (I.C.)
Manual Import (from I.C.)
Manual Select Strand & Oligo(s)
SNPOligo
W3_SNPSet
OligoQueue & W3_SNPOligo
2. Synthesis
Insert
Insert
Insert
Insert
Insert
Insert

Figure 58c

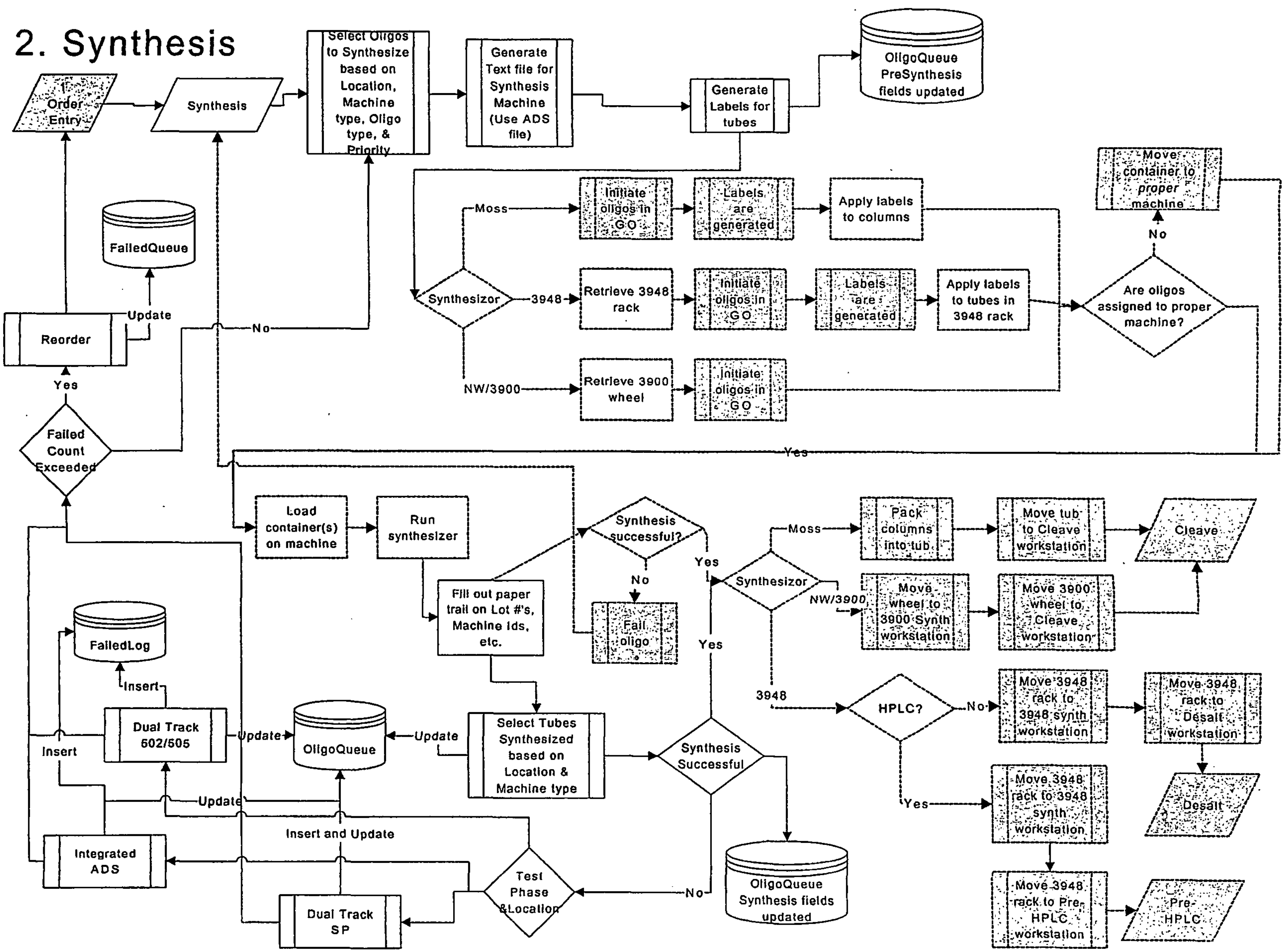
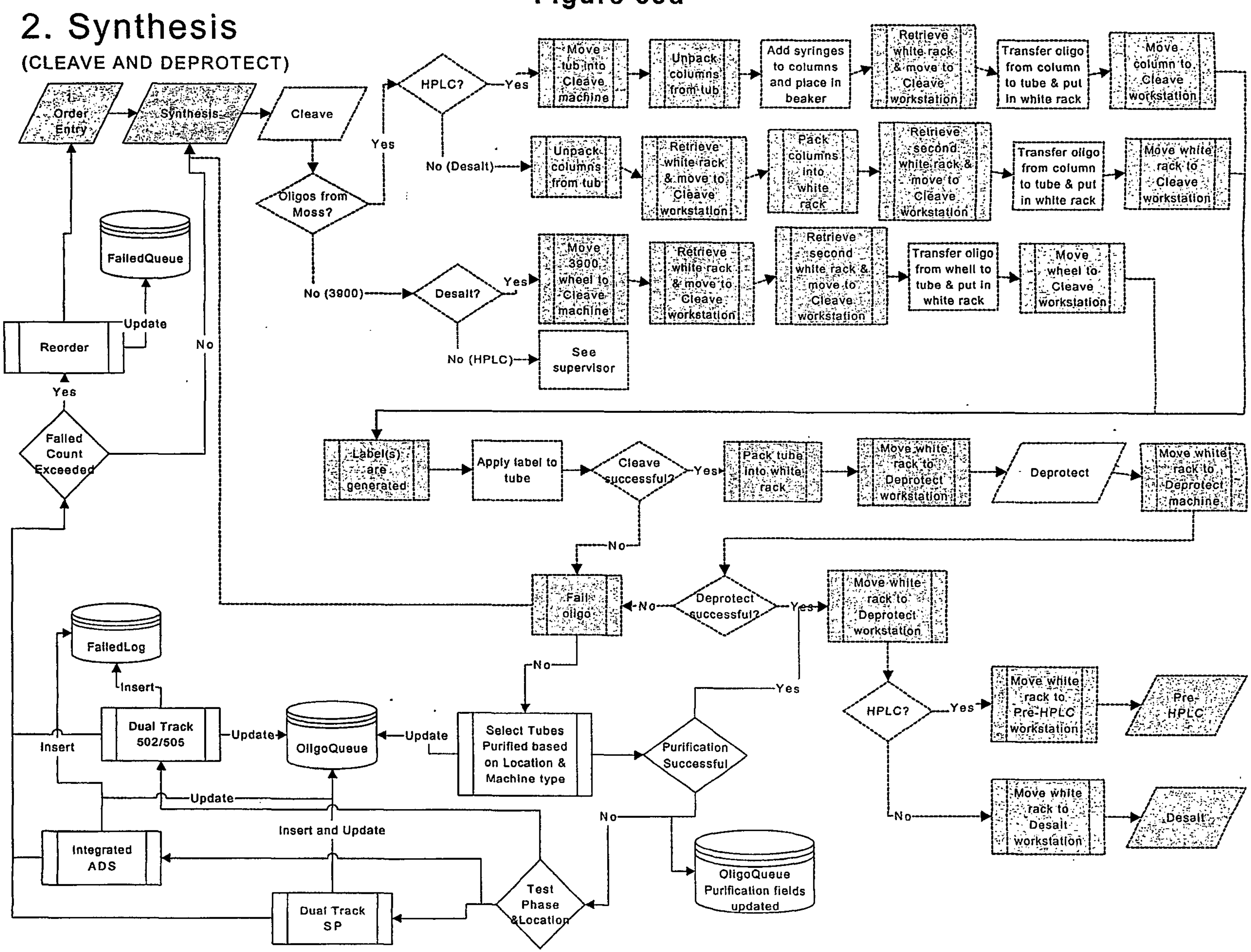

2. Synthesis
Order Entry
Synthesis
Select Oligos to Synthesize based on Location, Machine type, Oligo type, & Priority
Generate Text file for Synthesis Machine (Use ADS file)
Generate Labels for tubes
OligoQueue PreSynthesis fields updated
FailedQueue
Update
Reorder
Yes
No
Failed Count Exceeded
Synthesizor
Moss
3948
NW/3900
Initiate oligos in GO
Labels are generated
Apply labels to columns
Retrieve 3948 rack
Initiate oligos in GO
Labels are generated
Apply labels to tubes in 3948 rack
Retrieve 3900 wheel
Initiate oligos in GO
Are oligos assigned to proper machine?
No
Move container to proper machine
Yes
Load container(s) on machine
Run synthesizer
Fill out paper trail on Lot #'s, Machine Ids, etc.
Synthesis successful?
No
Fail oligo
Yes
Synthesizor
Moss
NW/3900
3948
Pack columns into tub
Move tub to Cleave workstation
Cleave
Move wheel to 3900 Synth workstation
Move 3900 wheel to Cleave workstation
HPLC?
No
Yes
Move 3948 rack to 3948 synth workstation
Move 3948 rack to Desalt workstation
Desalt
Move 3948 rack to 3948 synth workstation
Move 3948 rack to Pre-HPLC workstation
Pre-HPLC
FailedLog
Insert
Dual Track 502/505
Update
OligoQueue
Update
Select Tubes Synthesized based on Location & Machine type
Synthesis Successful
Yes
No
OligoQueue Synthesis fields updated
Insert
Update
Insert and Update
Integrated ADS
Test Phase &Location
Dual Track SP

Figure 58d


2. Synthesis
(CLEAVE AND DEPROTECT)
Order Entry
Synthesis
Cleave
Oligos from Moss?
Yes
HPLC?
Yes
Move tub into Cleave machine
Unpack columns from tub
Add syringes to columns and place in beaker
Retrieve white rack & move to Cleave workstation
Transfer oligo from column to tube & put in white rack
Move column to Cleave workstation
No (Desalt)
Unpack columns from tub
Retrieve white rack & move to Cleave workstation
Pack columns into white rack
Retrieve second white rack & move to Cleave workstation
Transfer oligo from column to tube & put in white rack
Move white rack to Cleave workstation
No (3900)
Desalt?
Yes
Move 3900 wheel to Cleave machine
Retrieve white rack & move to Cleave workstation
Retrieve second white rack & move to Cleave workstation
Transfer oligo from wheel to tube & put in white rack
Move wheel to Cleave workstation
No (HPLC)
See supervisor
FailedQueue
Update
Reorder
No
Yes
Failed Count Exceeded
Label(s) are generated
Apply label to tube
Cleave successful?
Yes
Pack tube into white rack
Move white rack to Deprotect workstation
Deprotect
Move white rack to Deprotect machine
No
Fail oligo
No
Deprotect successful?
Yes
Move white rack to Deprotect workstation
HPLC?
Yes
Move white rack to Pre-HPLC workstation
Pre-HPLC
No
Move white rack to Desalt workstation
Desalt
FailedLog
Insert
Dual Track 502/505
Update
OligoQueue
Update
Select Tubes Purified based on Location & Machine type
Purification Successful
Yes
No
OligoQueue Purification fields updated
Insert
Update
Insert and Update
Integrated ADS
Dual Track SP
Test Phase &Location

2. Synthesis
(Pre-HPLC and HPLC)

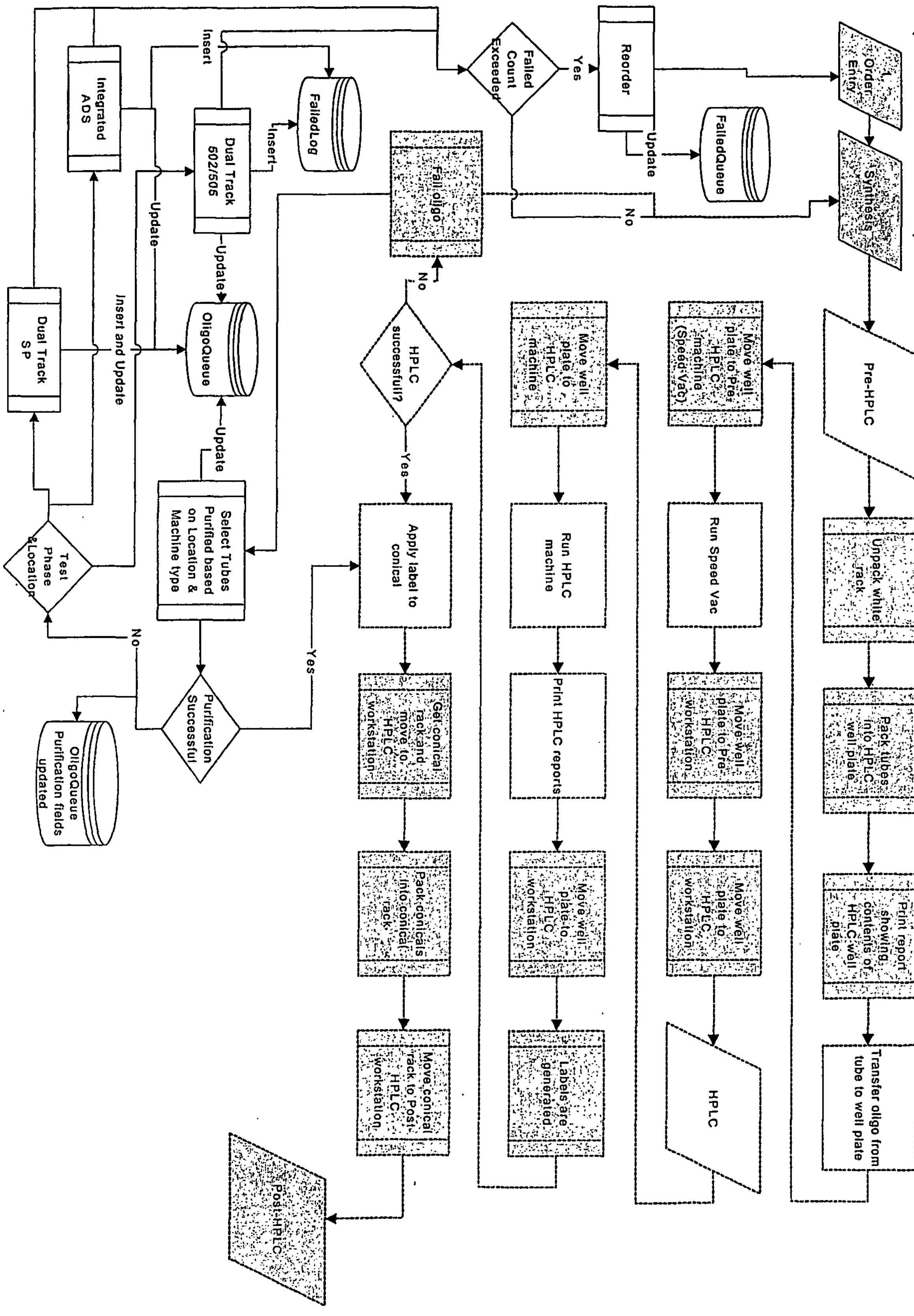
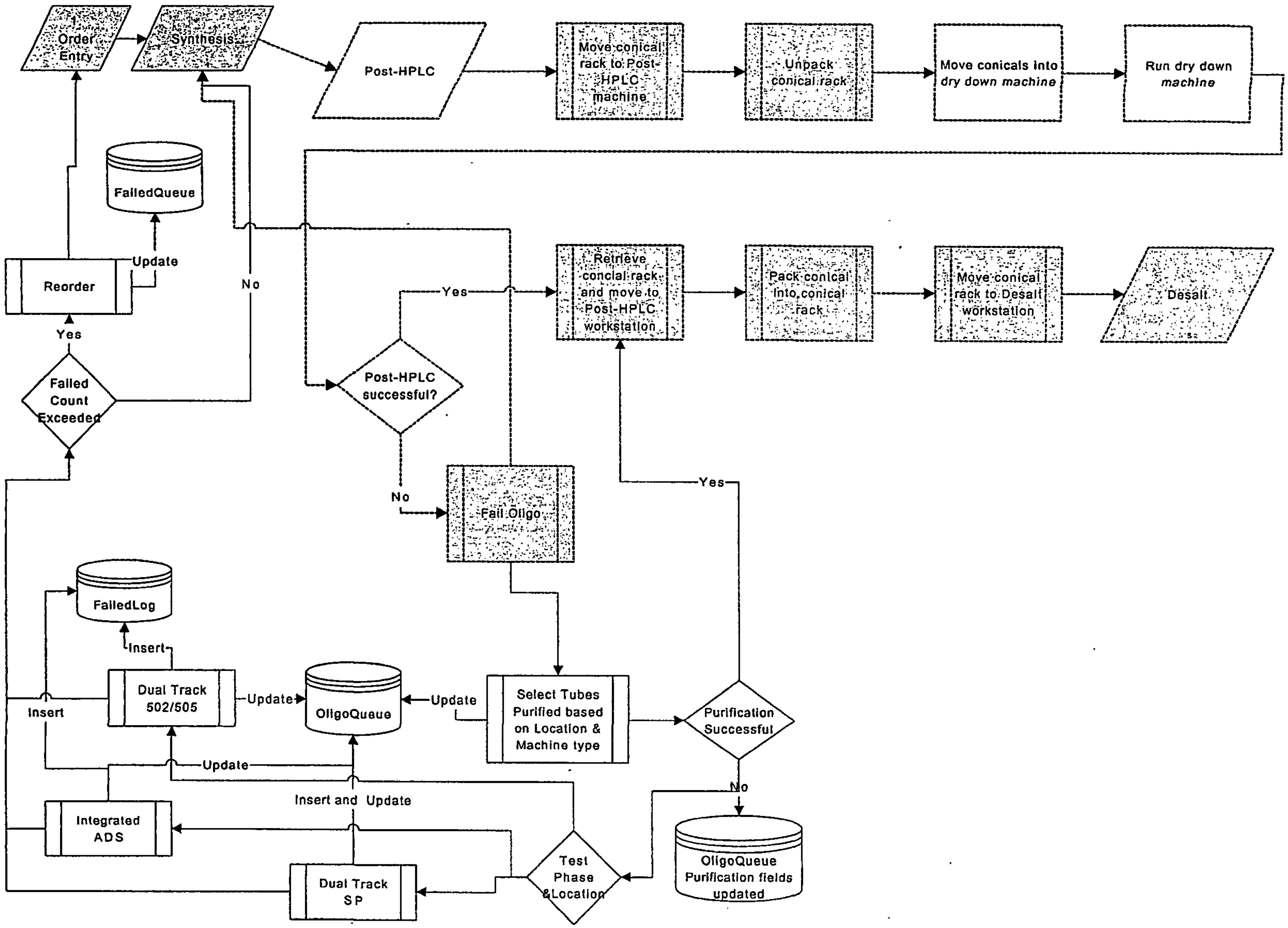


Figure 58e

Figure 58f

2. Synthesis

(Post-HPLC)



2. Synthesis
(Desalt)

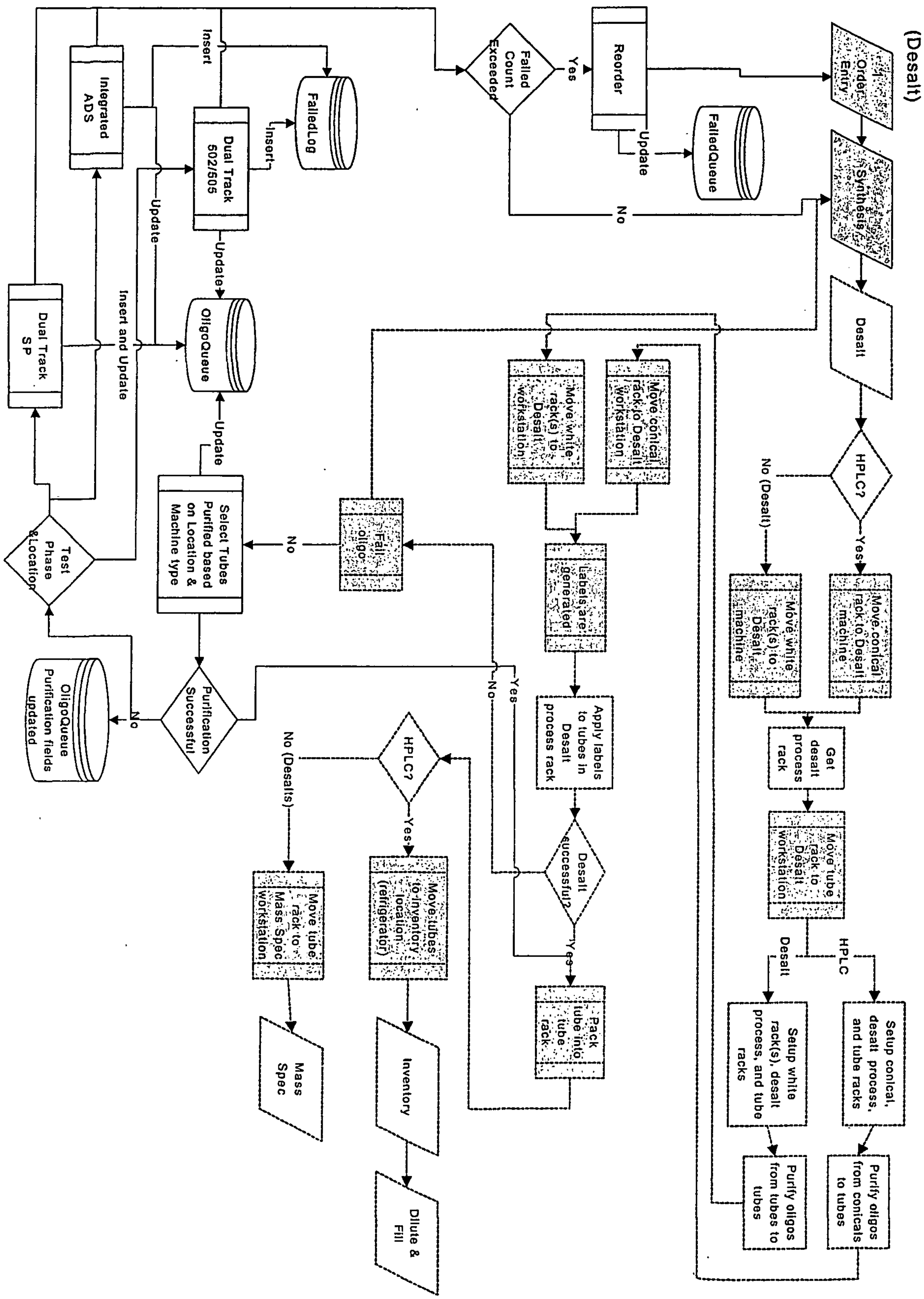
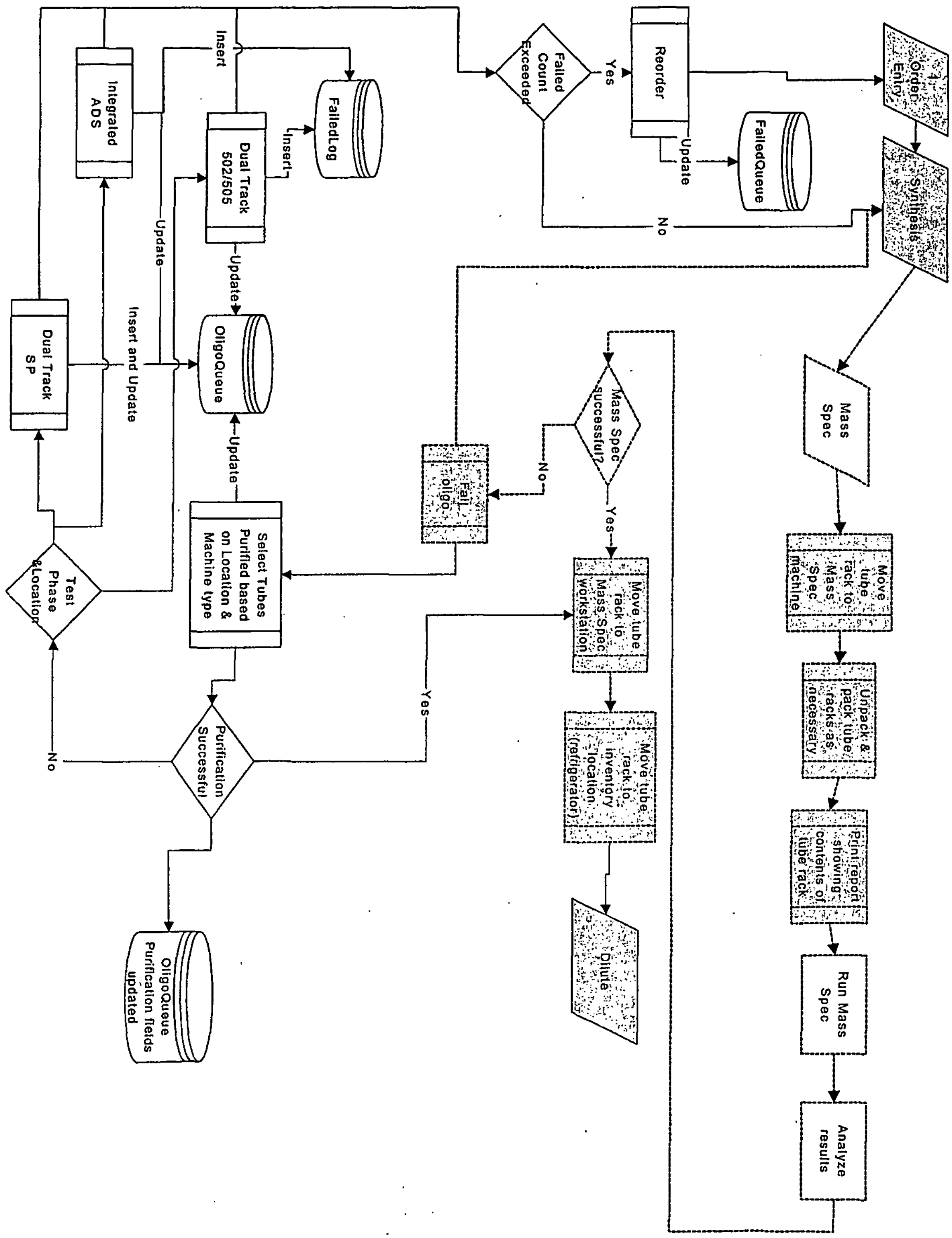


Figure 58g

Figure 58h

2. Synthesis
(Mass Spec)



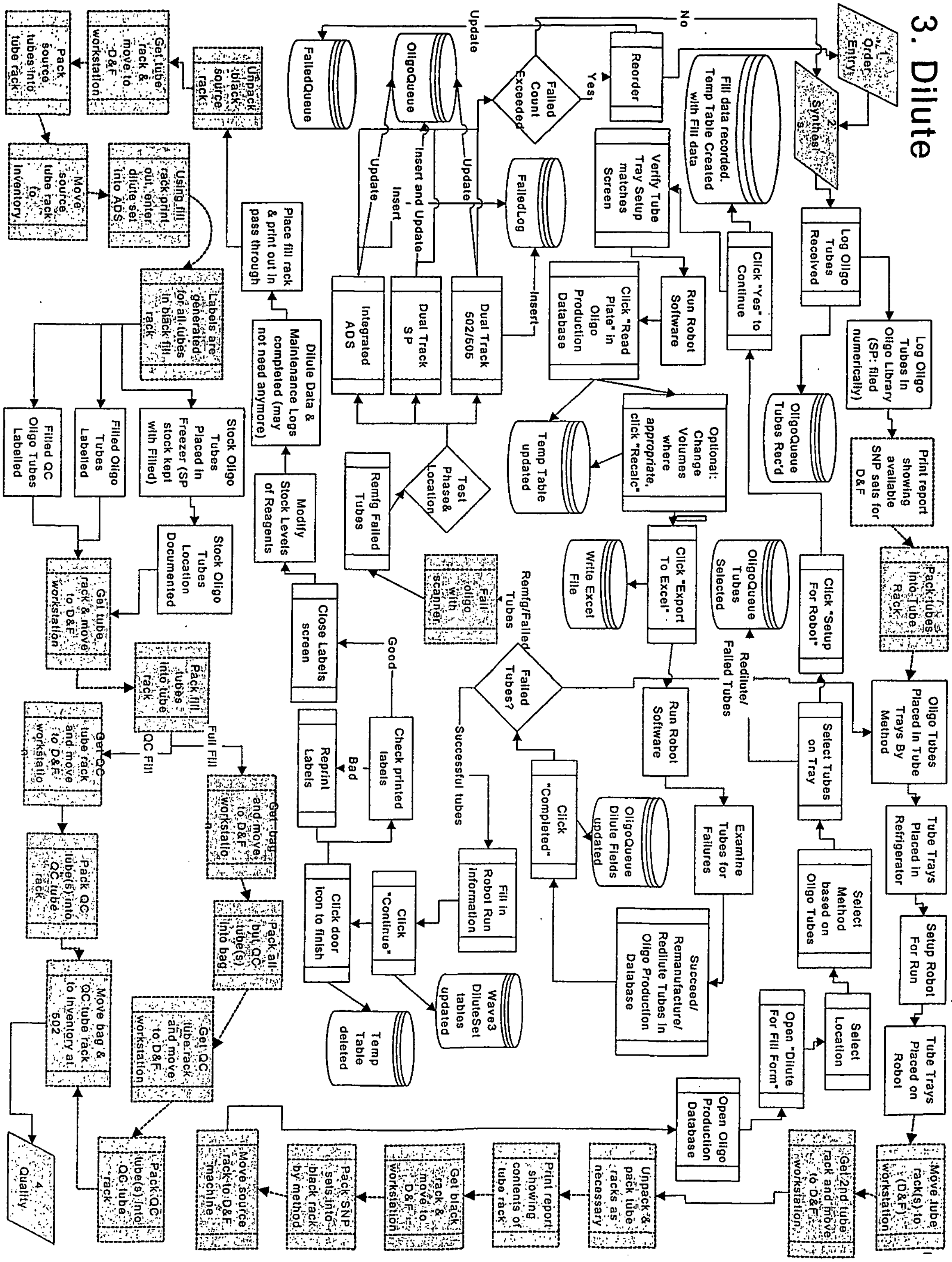

3. Dilute
1. Order Entry
2. Synthesis
Log Oligo Tubes Received
Log Oligo Tubes In Oligo Library (SP: filed numerically)
Print report showing available SNP sets for D&F
Pack tubes into Tube Rack
Oligo Tubes Placed in Tube Trays By Method
Tube Trays Placed in Refrigerator
Setup Robot For Run
Tube Trays Placed on Robot
Move tube rack(s) to (D&F) workstation
Get 2nd tube rack and move to D&F workstation
Unpack & pack tube racks as necessary
Print report showing contents of tube rack
Get black rack & move to D&F workstation
Pack SNP sets into black rack by method
Move source rack to D&F machine
Open Oligo Production Database
Open "Dilute For Fill Form"
Select Location
Select Method based on Oligo Tubes
Select Tubes on Tray
Click "Setup For Robot"
OligoQueue Tubes Rec'd
Click "Yes" to Continue
Fill data recorded. Temp Table Created with Fill data
Reorder
Yes
No
Failed Count Exceeded
Update
OligoQueue
FailedQueue
FailedLog
Verify Tube Tray Setup matches Screen
Run Robot Software
Click "Read Plate" in Oligo Production Database
Optional: Change Volumes where appropriate, click "Recalc"
Temp Table updated
Click "Export To Excel"
Write Excel File
OligoQueue Tubes Selected
Redilute/ Failed Tubes
Run Robot Software
Examine Tubes for Failures
Succeed/ Remanufacture/ Redilute Tubes In Oligo Production Database
Click "Completed"
OligoQueue Dilute Fields updated
Failed Tubes?
Remfg/Failed Tubes
Successful tubes
Fill in Robot Run Information
Click "Continue"
Wave3 DiluteSet tables updated
Click door icon to finish
Temp Table deleted
Fail oligo with scanner
Remfg Failed Tubes
Test Phase& Location
Dual Track 502/505
Dual Track SP
Integrated ADS
Insert
Insert and Update
Update
Dilute Data & Maintenance Logs completed (may not need anymore)
Modify Stock Levels of Reagents
Close Labels screen
Good
Check printed labels
Bad
Reprint Labels
Place fill rack & print out in pass through
Unpack black source rack
Get tube rack & move to D&F workstation
Pack source tubes into tube rack
Move source tube rack to Inventory
Using fill rack print out, enter dilute set into ADS
Labels are generated for all tubes in black fill rack
Stock Oligo Tubes Placed in Freezer (SP stock kept with Filled)
Filled Oligo Tubes Labelled
Filled QC Oligo Tubes Labelled
Stock Oligo Tubes Location Documented
Get tube rack & move to D&F workstation
Pack fill tubes into tube rack
Full Fill
QC Fill
Get QC tube rack and move to D&F workstation
Get bag and move to D&F workstation
Pack QC tube(s) into QC tube rack
Pack all but QC tube(s) into bag
Get QC tube rack and move to D&F workstation
Pack QC tube(s) into QC tube rack
Move bag & QC tube rack to inventory at 502
4. Quality

Figure 58i

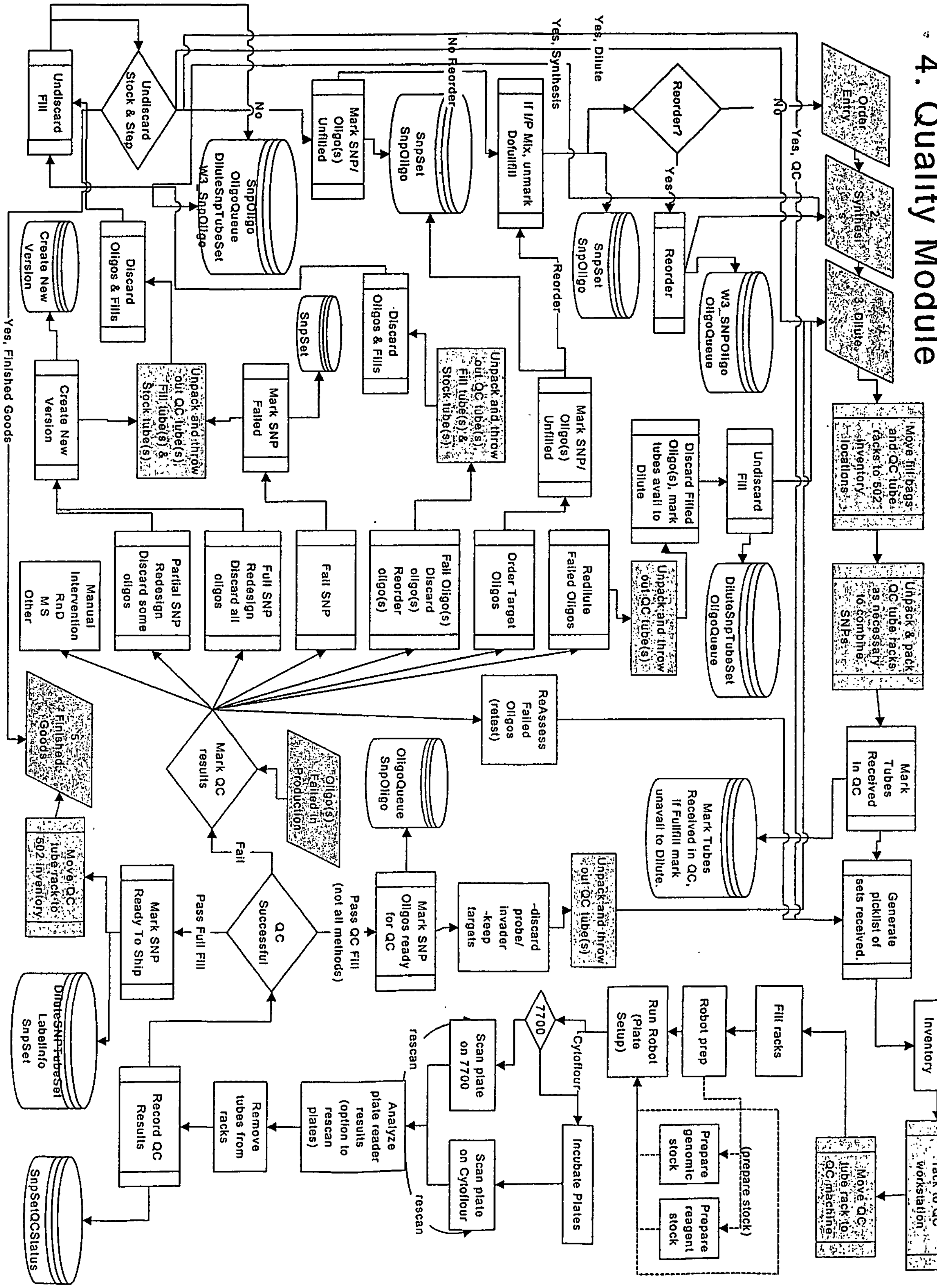

4. Quality Module
1. Order Entry
2. Synthesis
3. Dilute
Yes, QC
No
Reorder?
Yes
Reorder
W3_SNPOligo OligoQueue
SnpSet SnpOligo
If I/P Mix, unmark Dofullfill
Yes, Dilute
Yes, Synthesis
No Reorder
Reorder
Mark SNP/ Oligo(s) Unfilled
SnpSet SnpOligo
Undiscard Stock & Step
No
Undiscard Fill
SnpOligo OligoQueue DiluteSnpTubeSet W3_SnpOligo
Discard Oligos & Fills
Create New Version
Yes, Finished Goods
Discard Oligos & Fills
SnpSet
Unpack and throw out QC tube(s), Fill tube(s) & Stock tube(s)
Mark SNP Failed
Create New Version
Move fill bags and QC tube racks to 602 inventory locations
Undiscard Fill
Discard Filled Oligo(s), mark tubes avail to Dilute
DiluteSnpTubeSet OligoQueue
Unpack and throw out QC tube(s)
Redilute Failed Oligos
Order Target Oligos
Fail Oligo(s) Discard oligo(s) Reorder oligo(s)
Fail SNP
Full SNP Redesign Discard all oligos
Partial SNP Redesign Discard some oligos
Manual Intervention RnD MS Other
Unpack & pack QC tube racks as necessary to combine SNPs
ReAssess Failed Oligos (retest)
Mark Tubes Received in QC
Mark Tubes Received in QC, If Fullfill mark unavail to Dilute.
OligoQueue SnpOligo
Oligo(s) Failed in Production
Mark QC results
5. Finished Goods
Fail
Generate picklist of sets received.
Unpack and throw out QC tube(s)
-discard probe/ invader -keep targets
Mark SNP Oligos ready for QC
Pass QC Fill (not all methods)
QC Successful
Pass Full Fill
Mark SNP Ready To Ship
Move QC tube rack to 602 inventory
DiluteSNPTubeSet LabelInfo SnpSet
Pull Inventory
Fill racks
Robot prep
Run Robot (Plate Setup)
7700
Cytoflour
Scan plate on 7700
rescan
Incubate Plates
Scan plate on Cytoflour
rescan
Analyze plate reader results (option to rescan plates)
Remove tubes from racks
Record QC Results
SnpSetQCStatus
(prepare stock)
Prepare genomic stock
Prepare reagent stock
Move QC tube rack to QC workstation
Move QC tube rack to QC machine

Figure 58j

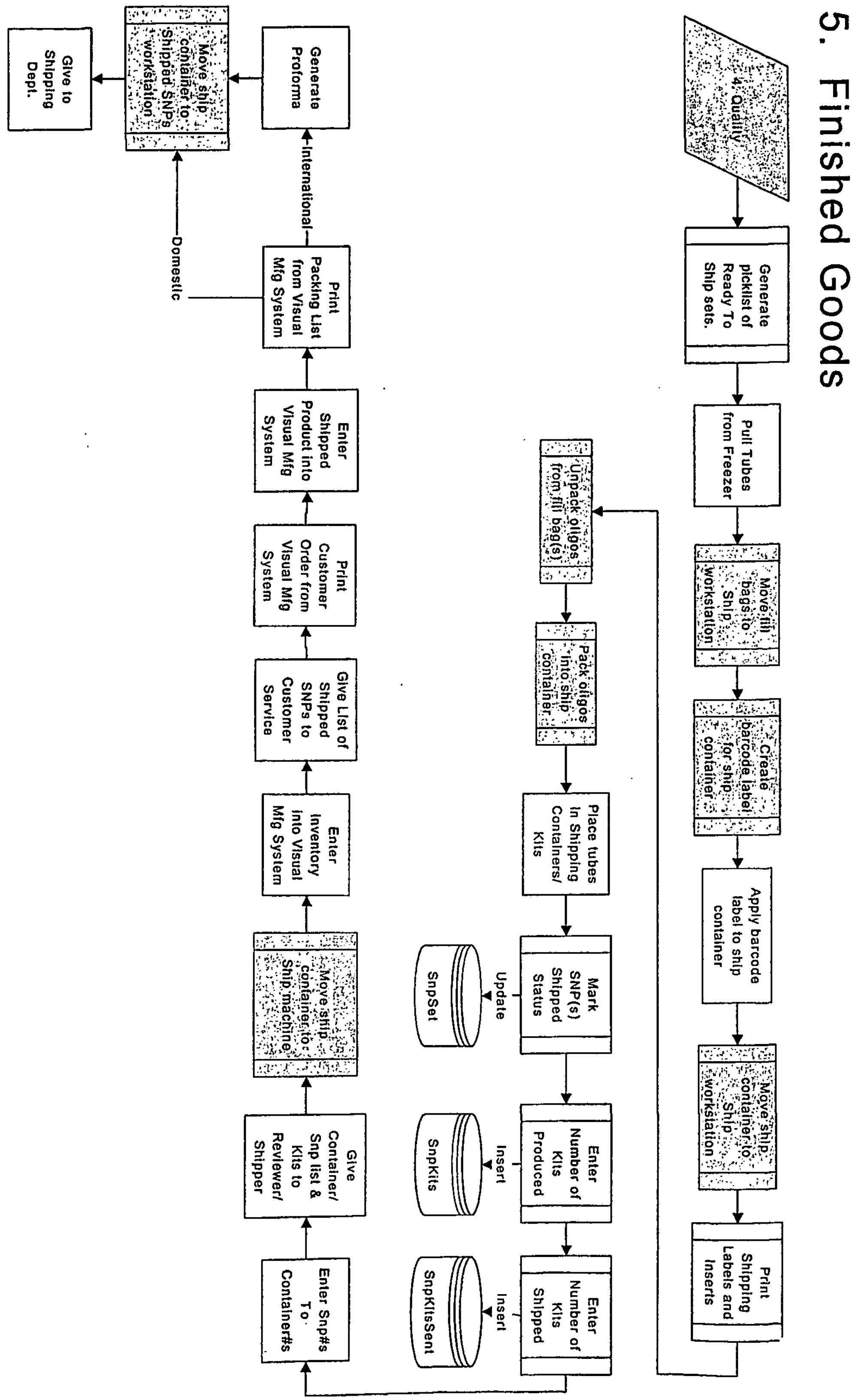

5. Finished Goods
4. Quality
Generate picklist of Ready To Ship sets.
Pull Tubes from Freezer
Move fill bags to Ship workstation
Create barcode label for ship container
Apply barcode label to ship container
Move ship container to Ship workstation
Print Shipping Labels and Inserts
Unpack oligos from fill bag(s)
Pack oligos into ship container
Place tubes in Shipping Containers/ Kits
Mark SNP(s) Shipped Status
Update
SnpSet
Enter Number of Kits Produced
Insert
SnpKits
Enter Number of Kits Shipped
Insert
SnpKitsSent
Enter Snp#s To Container#s
Give Container/ Snp list & Kits to Reviewer/ Shipper
Move ship container to Ship machine
Enter inventory into Visual Mfg System
Give List of Shipped SNPs to Customer Service
Print Customer Order from Visual Mfg System
Enter Shipped Product into Visual Mfg System
Print Packing List from Visual Mfg System
International
Domestic
Generate Proforma
Move ship container to Shipped SNPs workstation
Give to Shipping Dept.

Figure 58k

FIGURE 59A

Protocol for set-up 96 and/or 384-well Invader assays using the Biomek® 2000 CORE system.

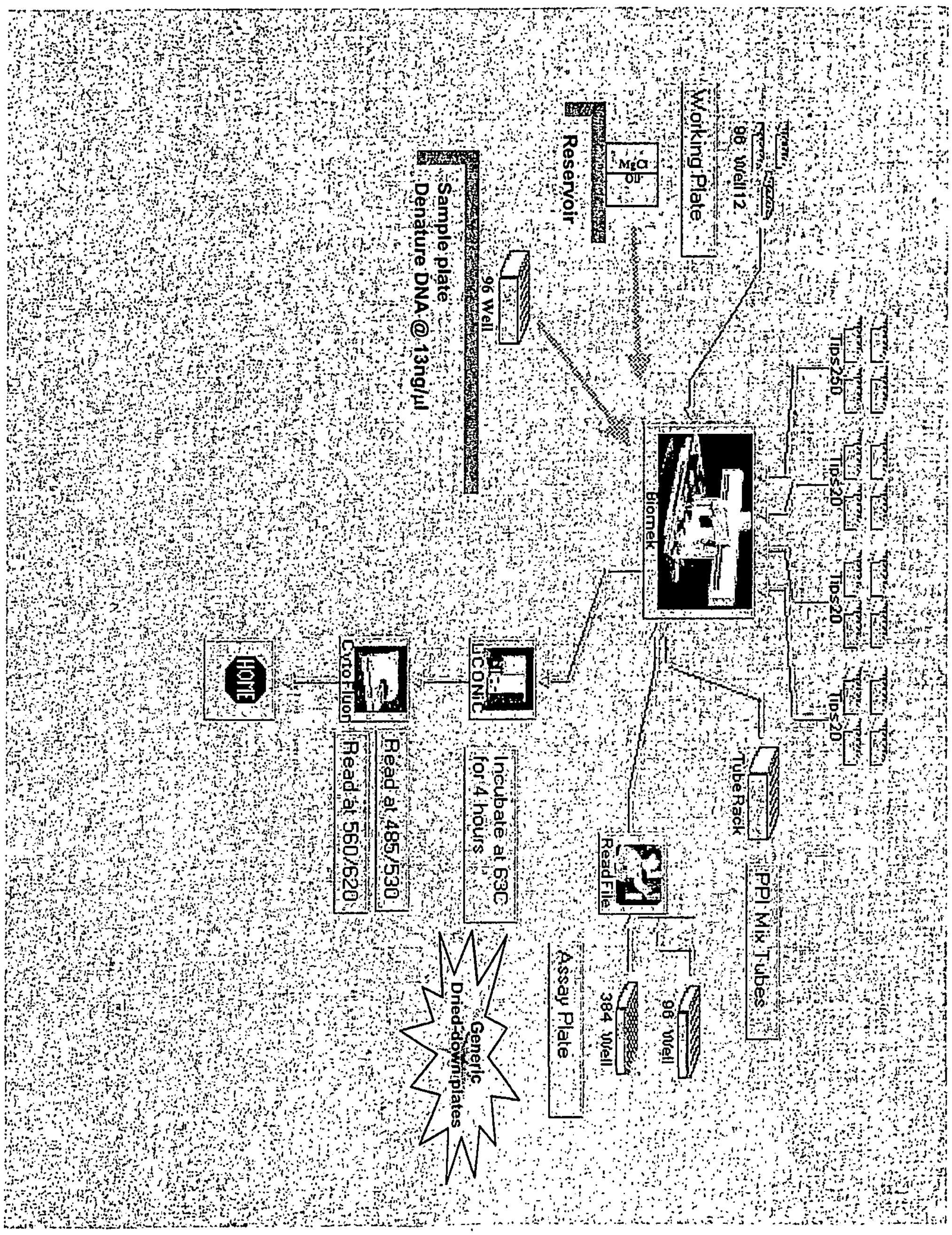


FIGURE 59B

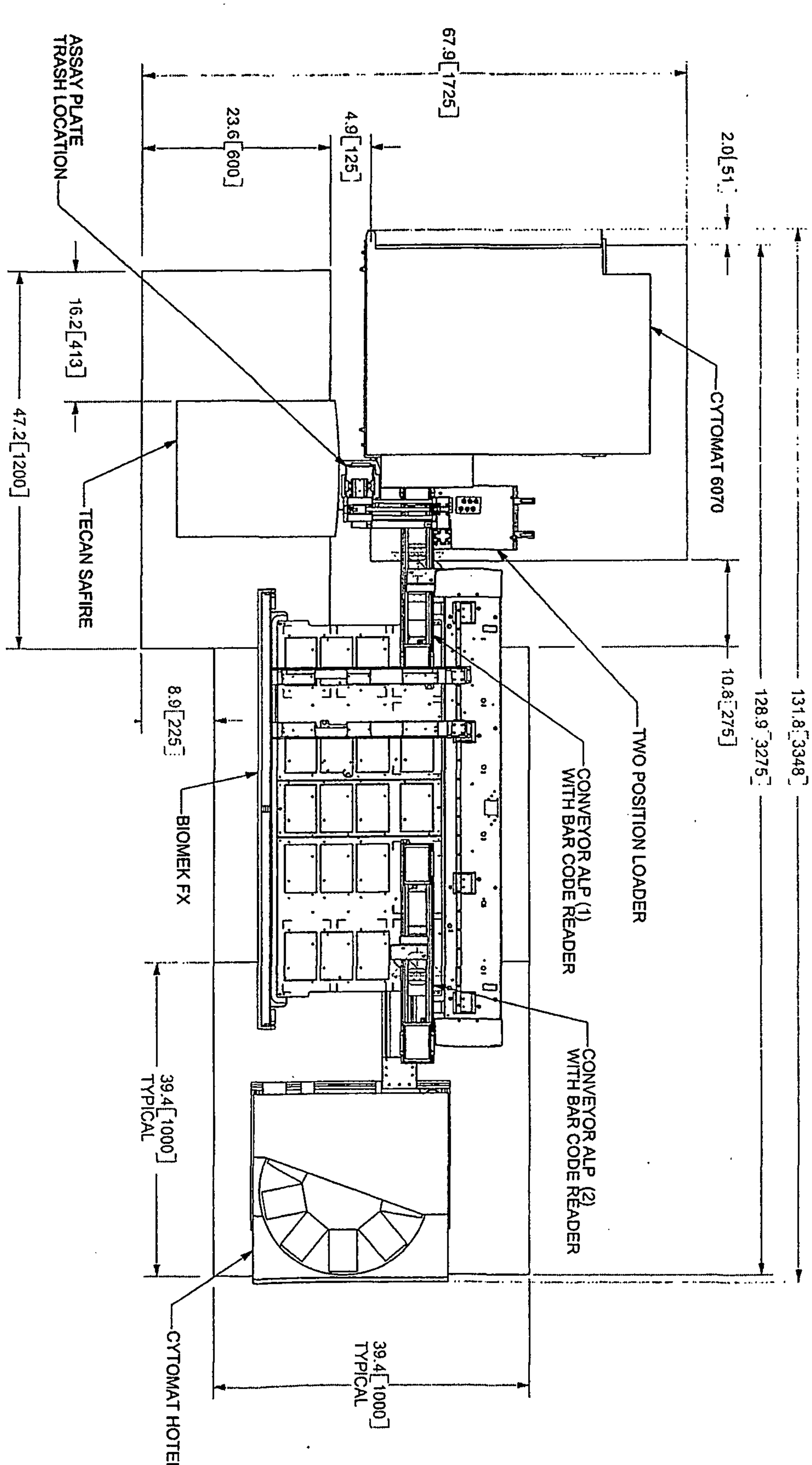
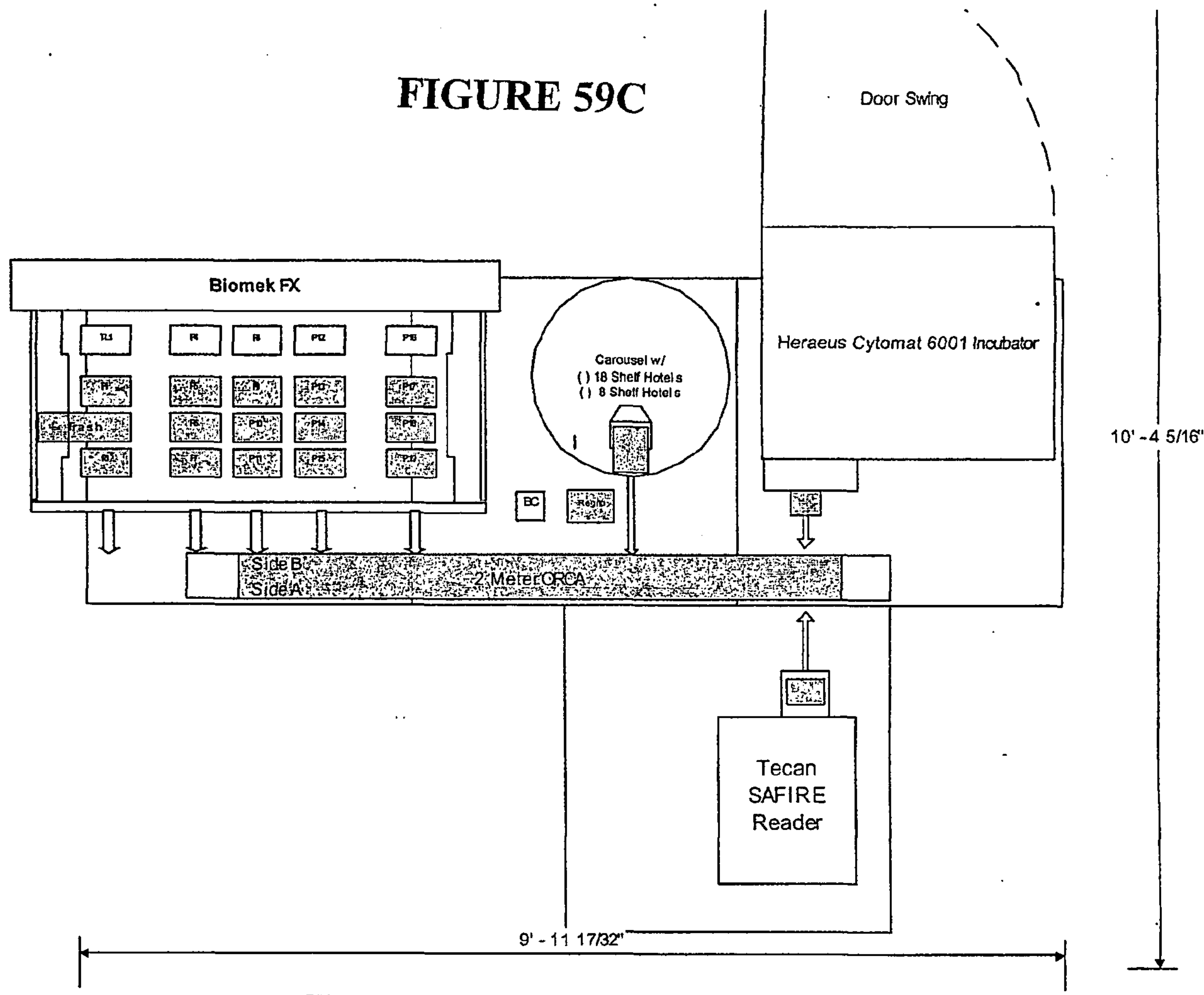

67.9 [1725]
23.6 [600]
4.9 [125]
2.0 [51]
ASSAY PLATE TRASH LOCATION
16.2 [413]
47.2 [1200]
CYTOMAT 6070
TECAN SAFIRE
TWO POSITION LOADER
CONVEYOR ALP (1) WITH BAR CODE READER
10.8 [275]
128.9 [3275]
131.8 [3348]
8.9 [225]
BIOMEK FX
CONVEYOR ALP (2) WITH BAR CODE READER
39.4 [1000] TYPICAL
CYTOMAT HOTEL
39.4 [1000] TYPICAL

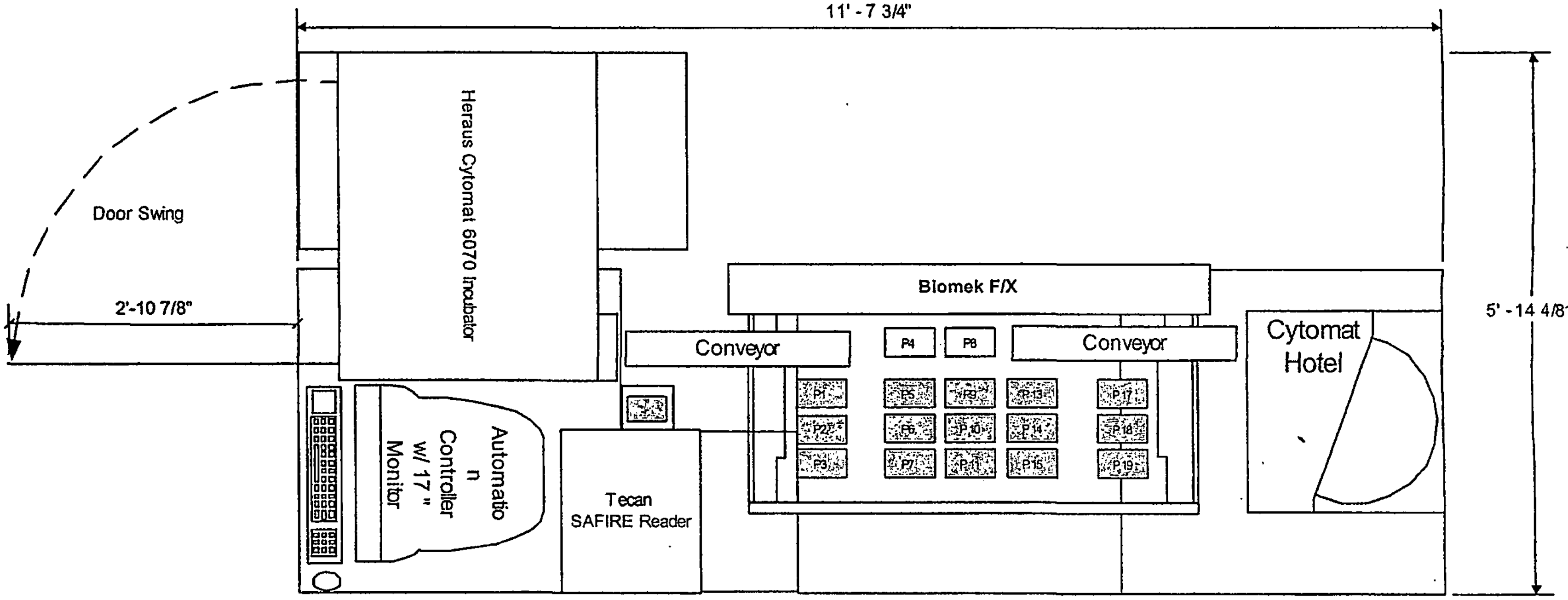
FIGURE 59C



Utilities
Electric
Air
Vacuum

-**Air Requirements:** 60 psi min.; 80 psi min for MP96 Pipettor. Utility Post terminations should end with a 1/4" NPT female fitting.
-**Vacuum Requirements:** 15-25 inHg, free air capacity of 75 l/min. Utility Post terminations should end with a 1/4" NPT female fitting.
-**Electrical Requirements:** Four 15Amp circuits, or three 20Amp circuits. The customer needs to supply a 6-outlet power strip for each ordered table and have an outlet available near each table to plug this into. If the customer is not going to supply 6-outlet power strips, 20-24 outlets should be spaced evenly along the back of the system.
-**CO2 Incubator, Tetrad:** These devices require 200-240v, 50/60 Hz, and include a 3-wire, unterminated cord. The customer must provide the desired plug to mate with the lab receptacle at time of installation.

FIGURE 59D



-Air Requirements: 60 psi min.; 80 psi min for MP96 Pipettor. Utility Post terminations should end with a 1/4" NPT female fitting.
-Vacuum Requirements: 15-25 InHg, free air capacity of 75 l/min. Utility Post terminations should end with a 1/4" NPT female fitting.
-Electrical Requirements: Four 15Amp circuits, or three 20Amp circuits. The customer needs to supply a 6-outlet power strip for each ordered table and have an outlet available near each table to plug this into. If the customer is not going to supply 6-outlet power strips, 20-24 outlets should be spaced evenly along the back of the system.
-CO2 Incubator, Tetrad: These devices require 200-240v, 50/60 Hz, and include a 3-wire, unterminated cord. The customer must provide the desired plug to mate with the lab receptacle at time of installation.

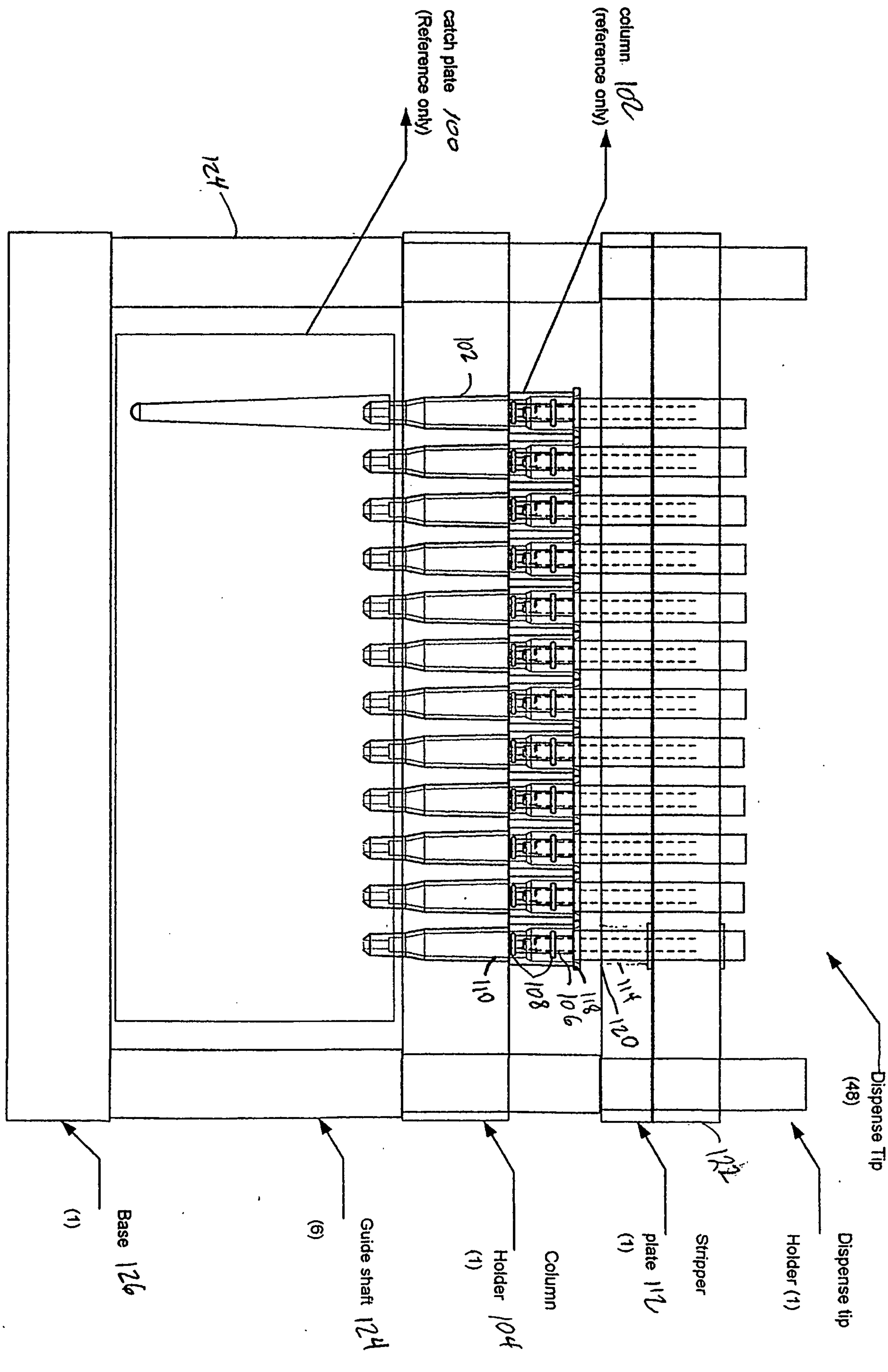

column 102
(reference only)
catch plate 100
(Reference only)
124
102
110
108
106
118
114
120
Dispense Tip
(48)
122
Dispense tip
Holder (1)
Stripper
plate 112
(1)
Column
Holder 104
(1)
Guide shaft 124
(6)
Base 126
(1)

FIGURE 60A

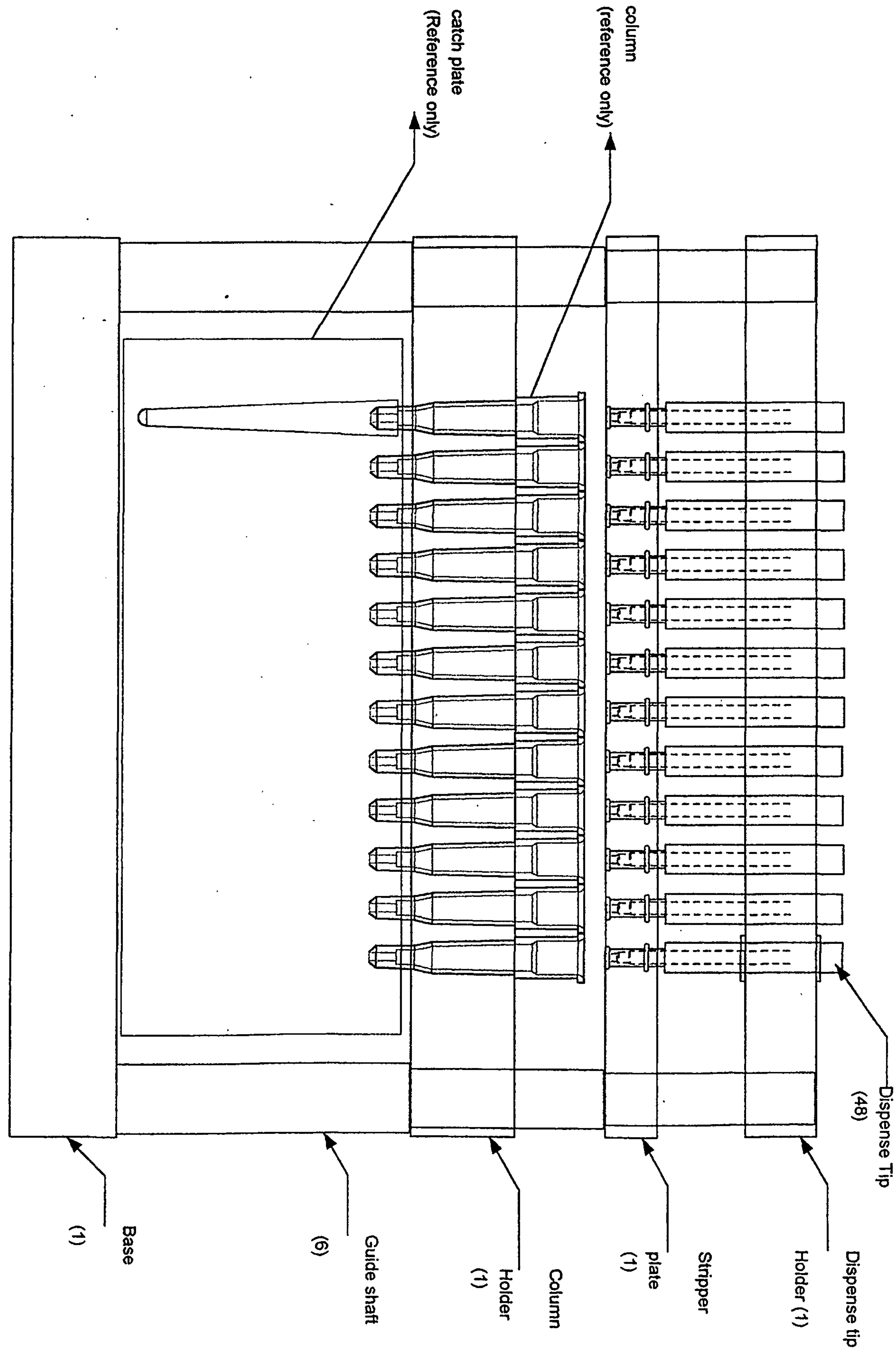

catch plate
(Reference only)
column
(reference only)
Dispense Tip
(48)
Base
(1)
Guide shaft
(6)
Column
Holder
(1)
Stripper
plate
(1)
Dispense tip
Holder (1)

FIGURE 60B

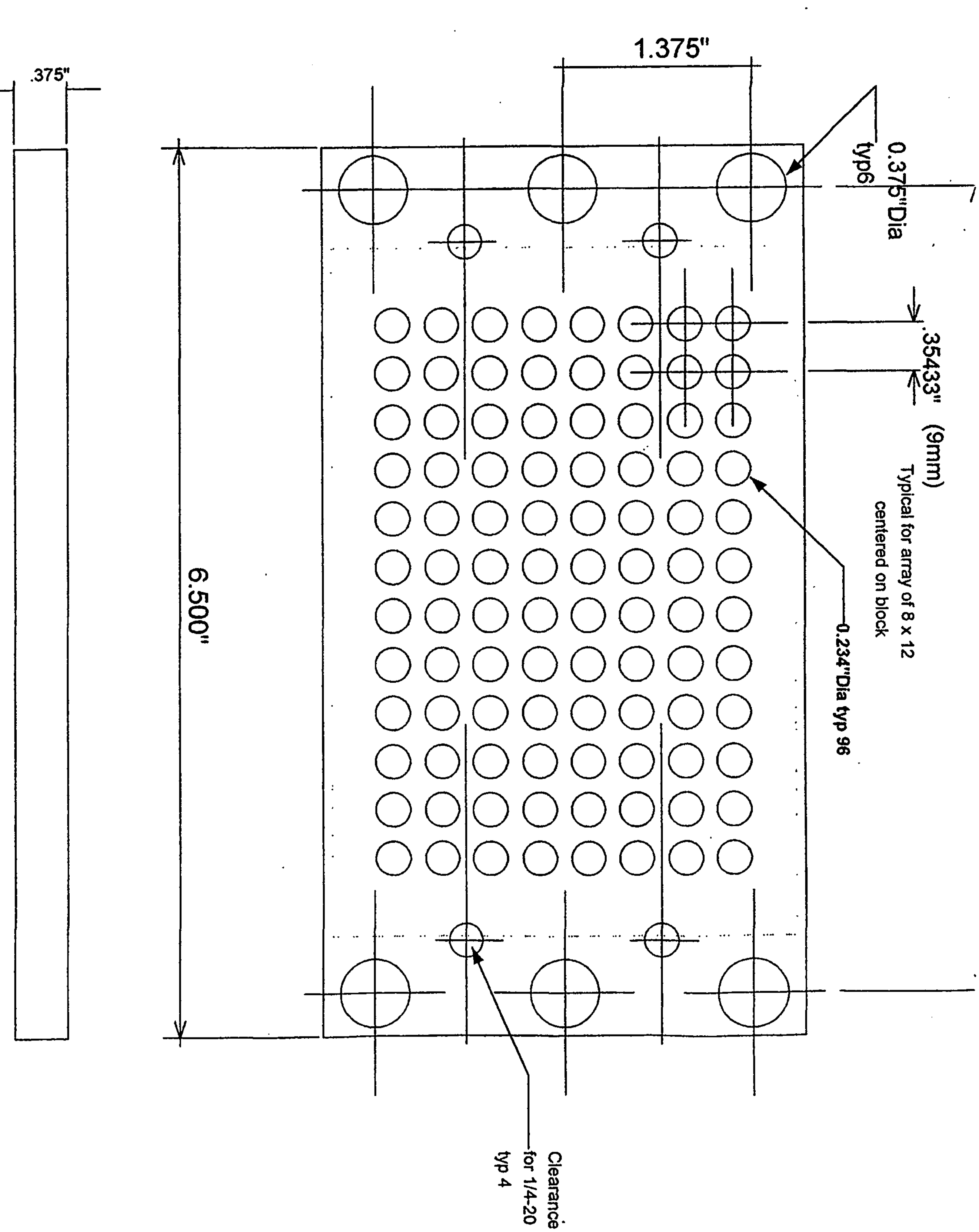

.375"
1.375"
0.375"Dia
typ6
.35433"
(9mm)
Typical for array of 8 x 12
centered on block
0.234"Dia typ 96
6.500"
Clearance
for 1/4-20
typ 4

FIGURE 60C

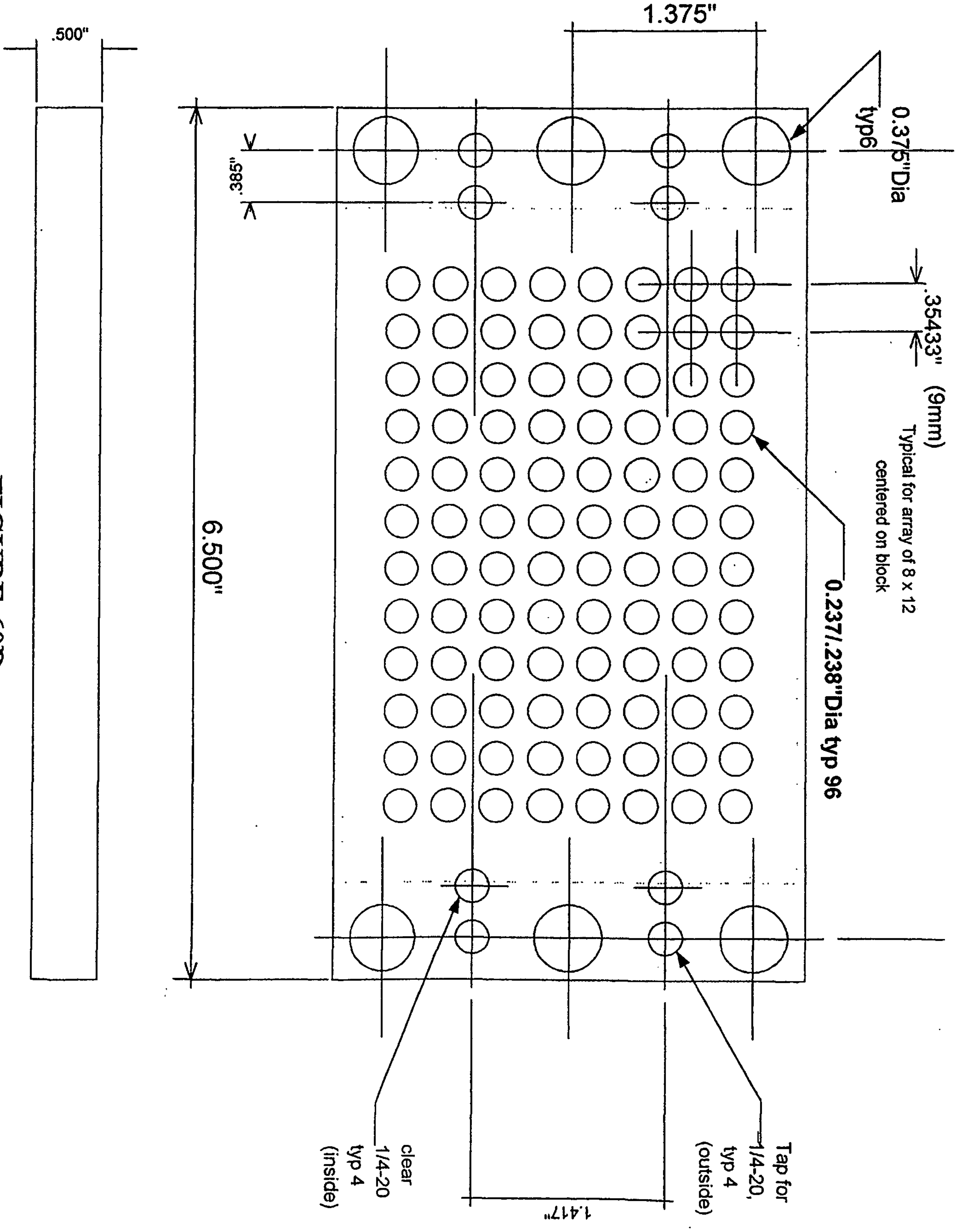

.500"
1.375"
0.375"Dia typ6
.385"
.35433" (9mm)
Typical for array of 8 x 12 centered on block
0.237/.238"Dia typ 96
6.500"
clear 1/4-20 typ 4 (inside)
Tap for 1/4-20, typ 4 (outside)
1.417"

FIGURE 60D

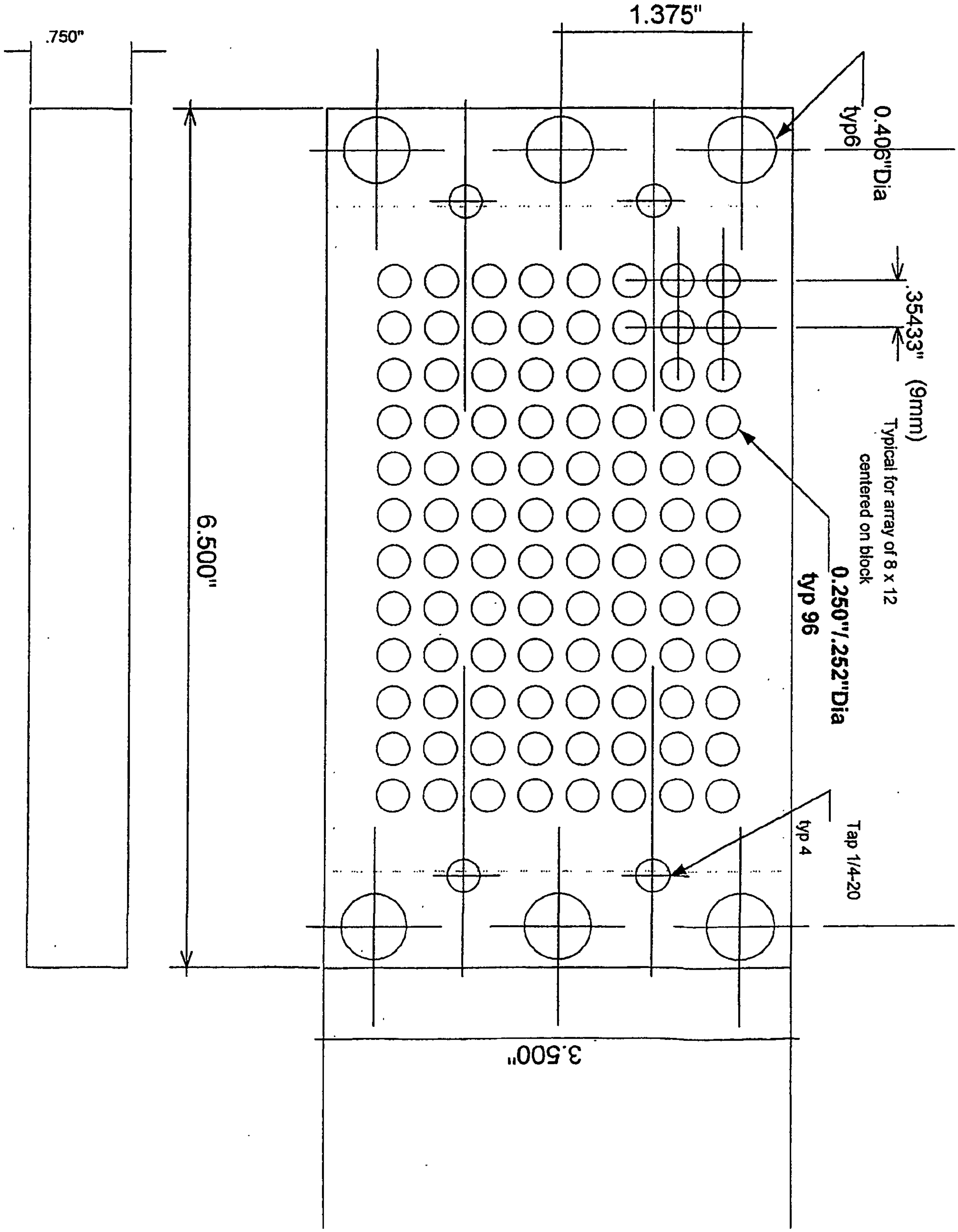

.750"
1.375"
0.406"Dia
typ6
.35433" (9mm)
Typical for array of 8 x 12
centered on block
0.250"/.252"Dia
typ 96
6.500"
Tap 1/4-20
typ 4
3.500"

FIGURE 60E

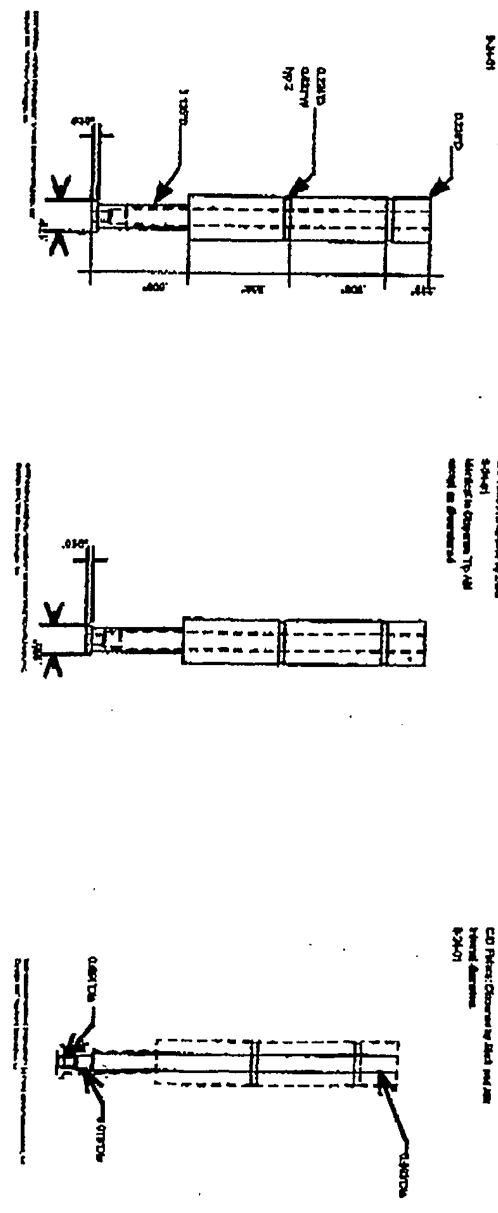


FIGURE 60F

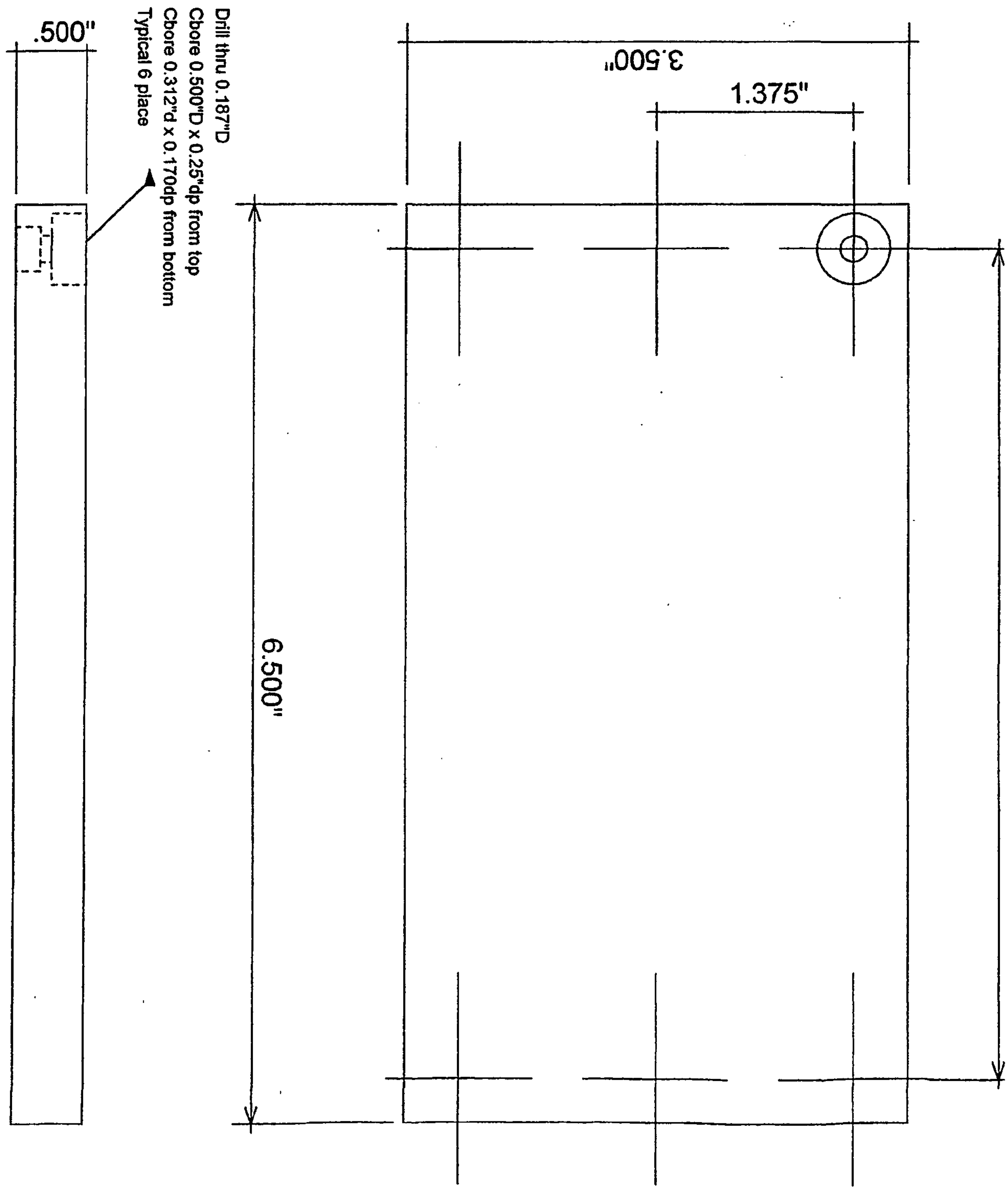

.500"
Drill thru 0.187"D
Cbore 0.500"D x 0.25"dp from top
Cbore 0.312"d x 0.170dp from bottom
Typical 6 place
3.500"
1.375"
6.500"

FIGURE 60G

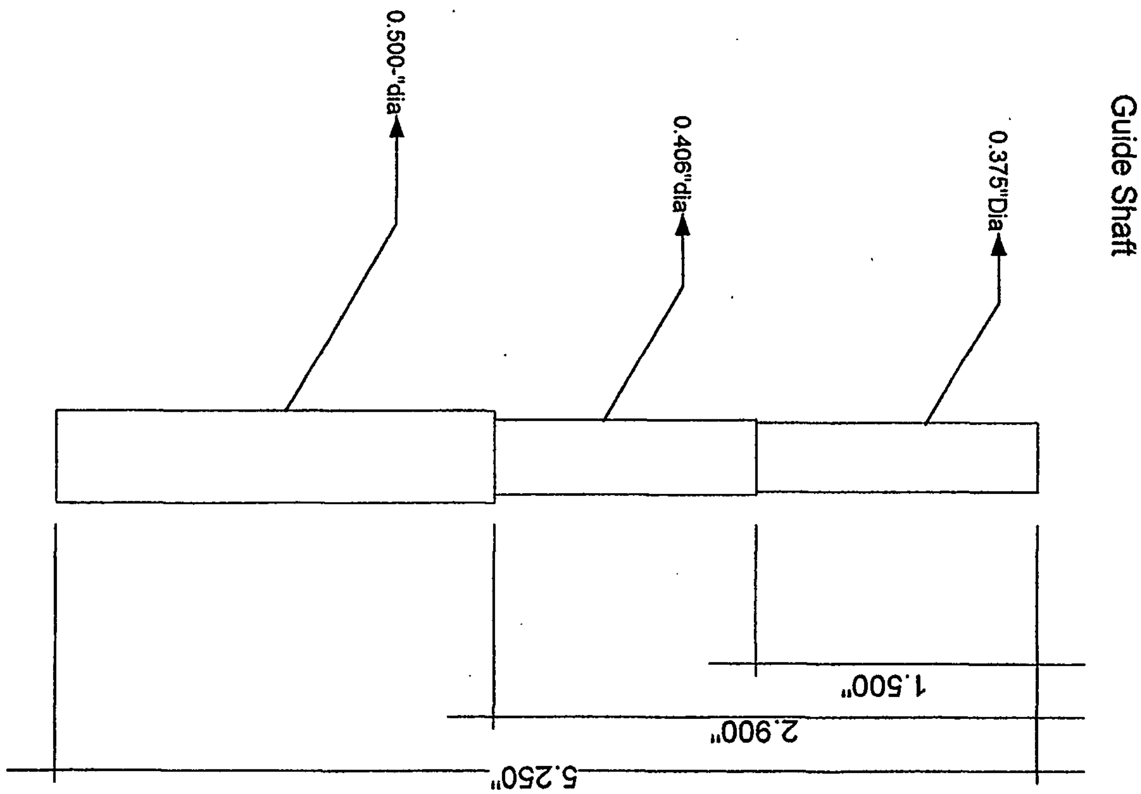


FIGURE 60H

FIGURE 61
DATA MANAGEMENT SYSTEM

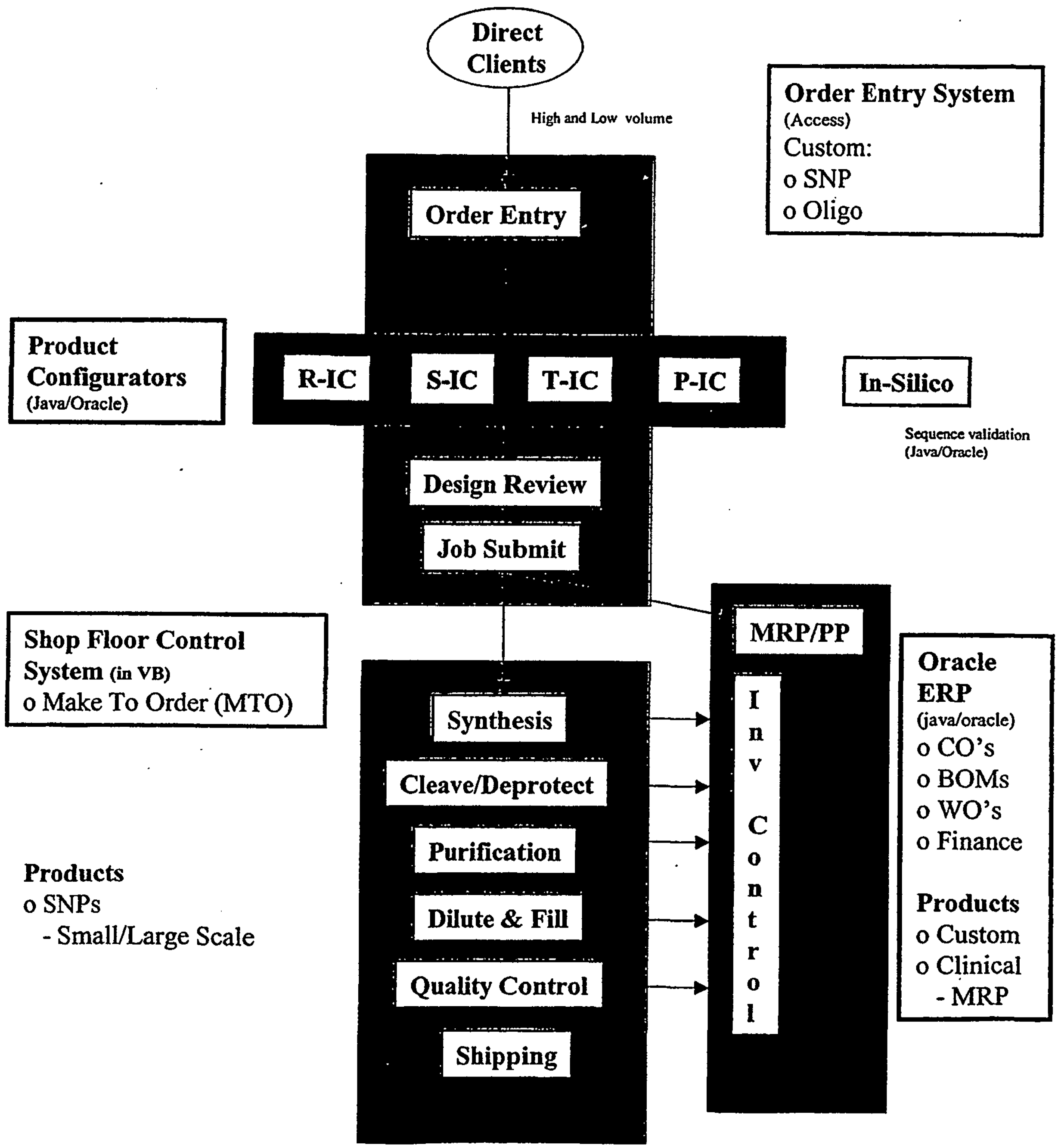


FIGURE 62
DATA MANAGEMENT SYSTEM

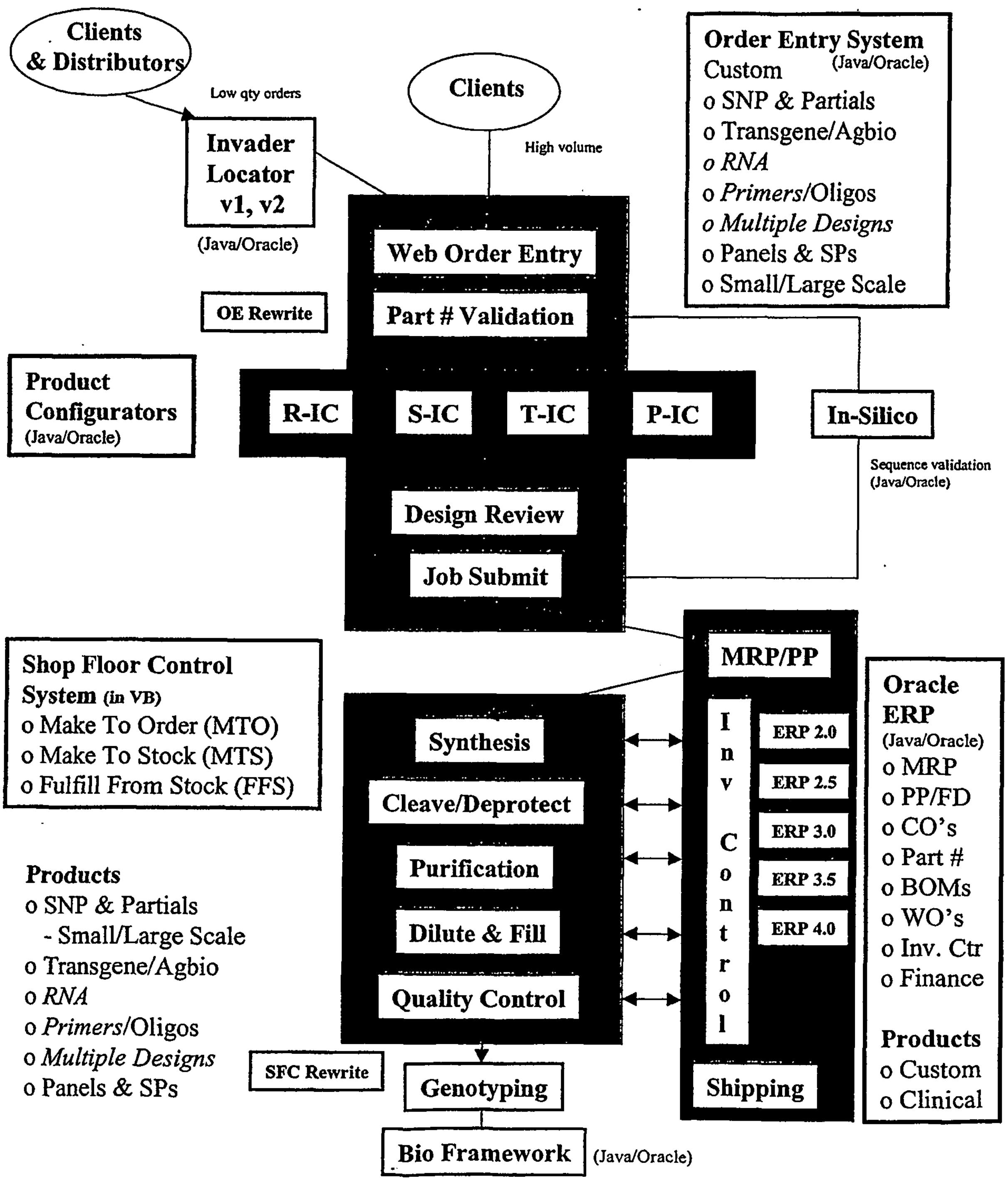


FIGURE 63

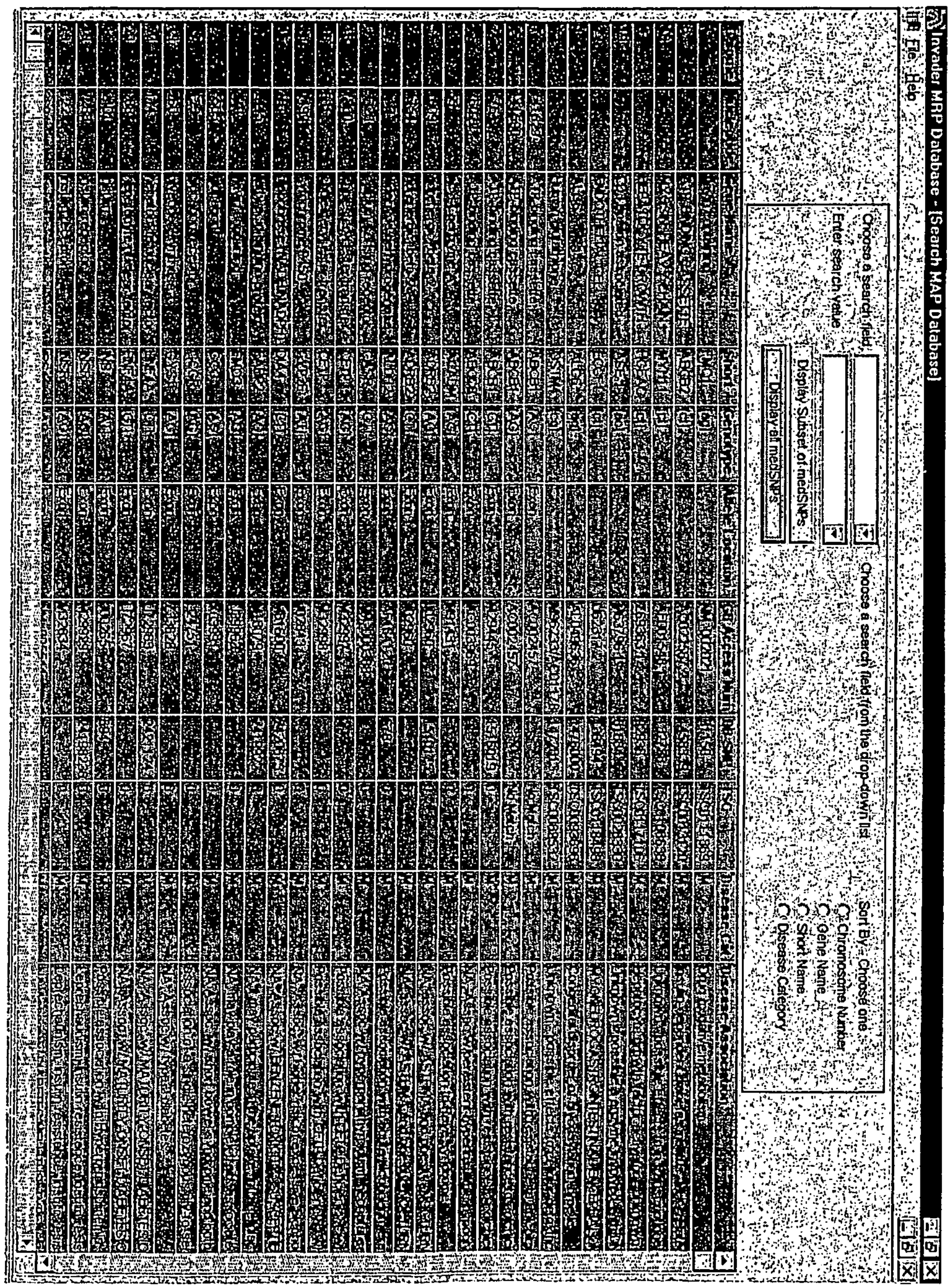

Invader MAP Database - [Search MAP Database]
File Help
Choose a search field
Enter search value
Display Subset of medSNPs
Display all medSNPs
Choose a search field from the drop-down list
Sort By - Choose one
Chromosome Number
Gene Name
Short Name
Disease Category

FIGURE 64

Microsoft Excel - Book2

A1 C'some

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	C'some	Nucleotid	Gene Nar	Short	Genotype	Allele Loc	Accession	DB SNP	TSC	Disease C	Disease A	OMIM Ref	Well
2	1		flavin conta	FMO1	[ag]		NM_00202	91150	TSC01113	M	Drug oxida	http://www	1
3	6		ATP-BIND	ABCB2	[ct]		AC022567	1458339	TSC00131	M	Insulin-dep	http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=170260	2
4	16		MYOSIN,	MYH11	[ct]		AF001548	132738	TSC01086	M	Unknown -	http://www	3
5	4		HEPATOC	HGFAC	[ct]		Z69363.1	92936	TSC01171	M	Hepatocyt	http://www	4
6	19		CD3 epsilo	ASE-1	[ac]		M63796.1	81884	TSC00283	M	Unknown -	http://www	5
7	22		ENDOTHE	ECGF1	[ct]		U62317.2	136494	TSC01164	M	MYONEUR	http://www	6
8	1		ALG6	ALG6	[ag]		AL049636	1476000	TSC00936	M	Congenital	http://www	7
9	1		Human glu	GSTM4	[cg]		M96233/A	1472813	TSC00895	M	Unknown -	http://www	8
10	7	141529	ATP-bindir	ABCB1	[AG]	Exon	AC002457		No Match	M	Possible a	http://www	9
11	7	141590	ATP-bindir	ABCB1	[AG]	Exon	AC002457		No Match	M	Possible a	http://www	10
12	7	176	ATP-bindir	ABCB1	[CT]	Exon	M29445	1516019	0	M	Possible a	http://www	11
13	11	76022	ATP-bindir	ABCC8	[CT]	Exon	AC003969		0	M	NA, Assoc	http://www	12
14	1	55	acyl-Coen	ACADM	[AG]	Exon	M91431		0	M	Possible a	http://www	13
15	5	1309	adrenergic	ADRB2	[AG]	Exon	J02960	1510175	0	M	NA, assoc	http://www	14
16	5	1754	adrenergic	ADRB2	[CT]	Exon	J02960		0	M	NA, assoc	http://www	15
17	12	113	aldehyde d	ALDH2	[AG]	Exon	AF164120	676	0	M	Possible a	http://www	16
18	3	3328	calcium-se	CASR	[GT]	Exon	NM_000388		0	M	NA Assoc	http://www	17
19	16	1300	cholestery	CETP	[AG]	Exon	M32997		0	M	Possible a	http://www	18
20	16	681	cholestery	CETP	[AG]	Exon	M32998		0	M	Possible a	http://www	19
21	10	2492	cytochrom	CYP2E	[AT]	Exon	J02843		0	M	NA, Assoc	http://www	20
22	16	191	diaphorase	DIA4	[CT]	Exon	M81600	2420673	0	M	NA, Assoc	http://www	21
23	4	2445	fatty acid b	FABP2	[AG]	Exon	M18079	2419923	0	M	NA It may	http://www	22
24	1	279	flavin conta	FMO3	[AG]	Exon	U39963		0	M	NA, may p	http://www	23
25	2	5505	glucagon	GCG	[AG]	Exon	X03991		0	M	NA Assoc	http://www	24
26	17	123	glucagon r	GCGR	[AG]	Exon	L24751		0	M	associatio	http://www	25
27	19	260	glycogen s	GYS1	[AG]	Exon	Z33628		0	M	NA, Assoc	http://www	26

Sheet1 / Sheet2 / Sheet3

Ready

FIGURE 65

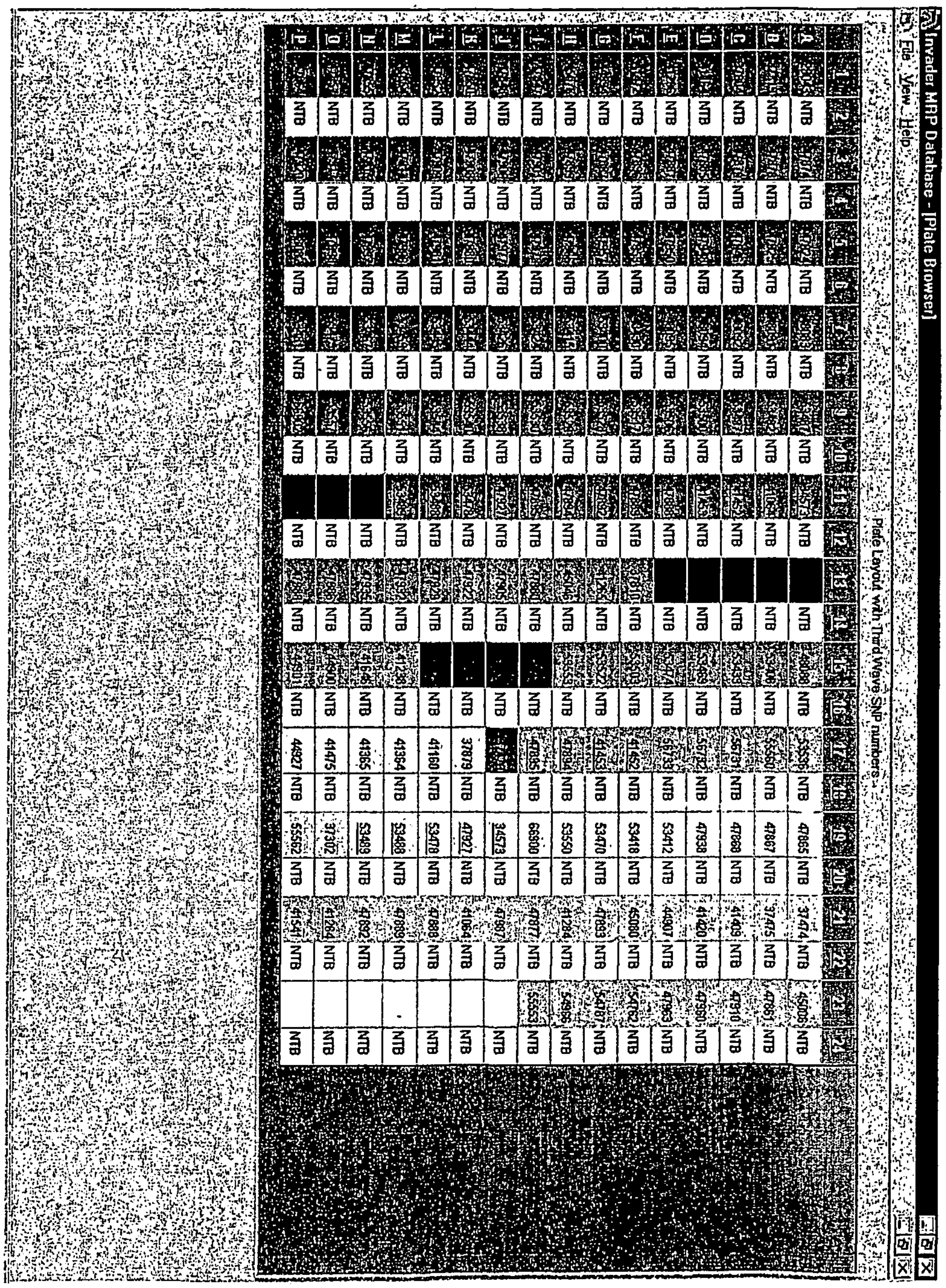
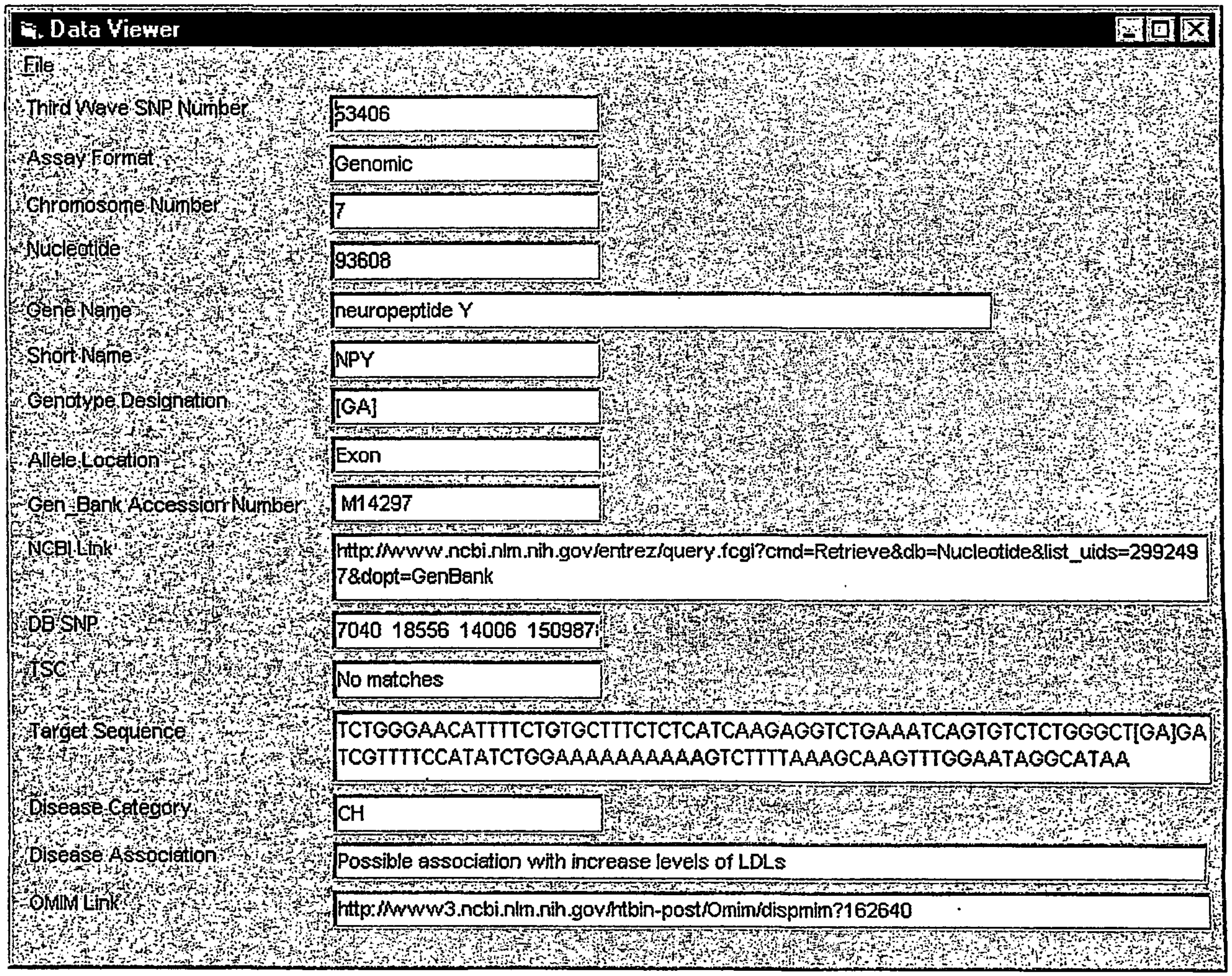

Invader MRP Database - [Plate Browser]
File View Help
Plate Layout with Third Wave SNP numbers

FIGURE 66



Data Viewer

File

Third Wave SNP Number	53406
Assay Format	Genomic
Chromosome Number	7
Nucleotide	93608
Gene Name	neuropeptide Y
Short Name	NPY
Genotype Designation	[GA]
Allele Location	Exon
Gen_Bank Accession Number	M14297
NCBI Link	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Nucleotide&list_uids=2992497&dopt=GenBank
DB SNP	7040 18556 14006 150987
TSC	No matches
Target Sequence	TCTGGGAACATTTTCTGTGCTTTCTCTCATCAAGAGGTCTGAAATCAGTGTCTCTGGGCT[GA]GATCGTTTTCCATATCTGGAAAAAAAAAAAGTCTTTTAAAGCAAGTTTGGAATAGGCATAA
Disease Category	CH
Disease Association	Possible association with increase levels of LDLs
OMIM Link	http://www3.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?162640

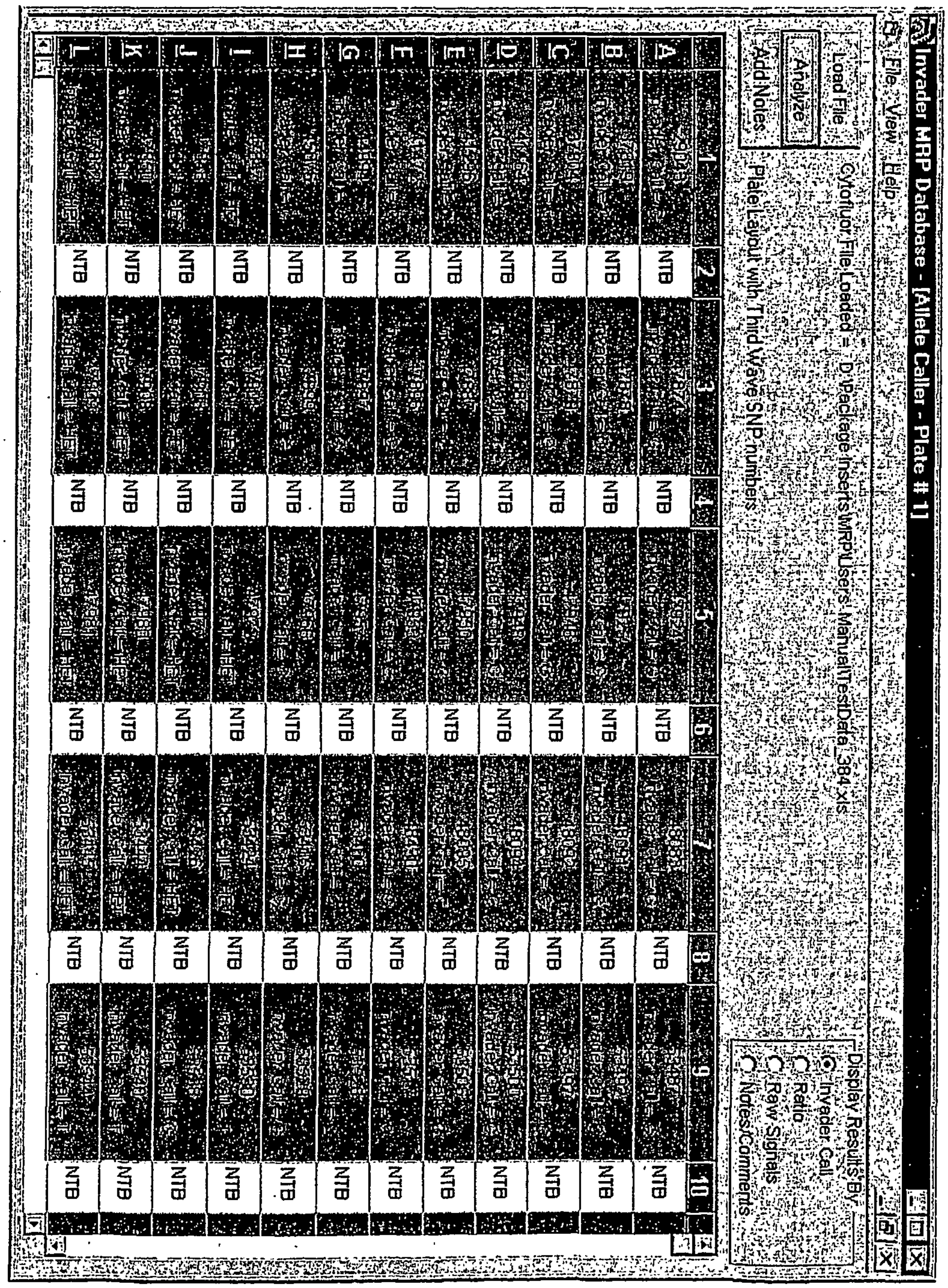

Invader MRP Database - [Allele Caller - Plate # 1]
File View Help
Load File
Analyze
Add Notes
Cytofluor File Loaded = D:\Package Inserts\MRP\Users Manual\TestData_384.xls
Plate Layout with Third Wave SNP numbers
Display Results By
Invader Call
Ratio
Raw Signals
Notes/Comments
1
2
3
4
5
6
7
8
9
10
A
B
C
D
E
F
G
H
I
J
K
L
NTB

FIGURE 67

FIGURE 68

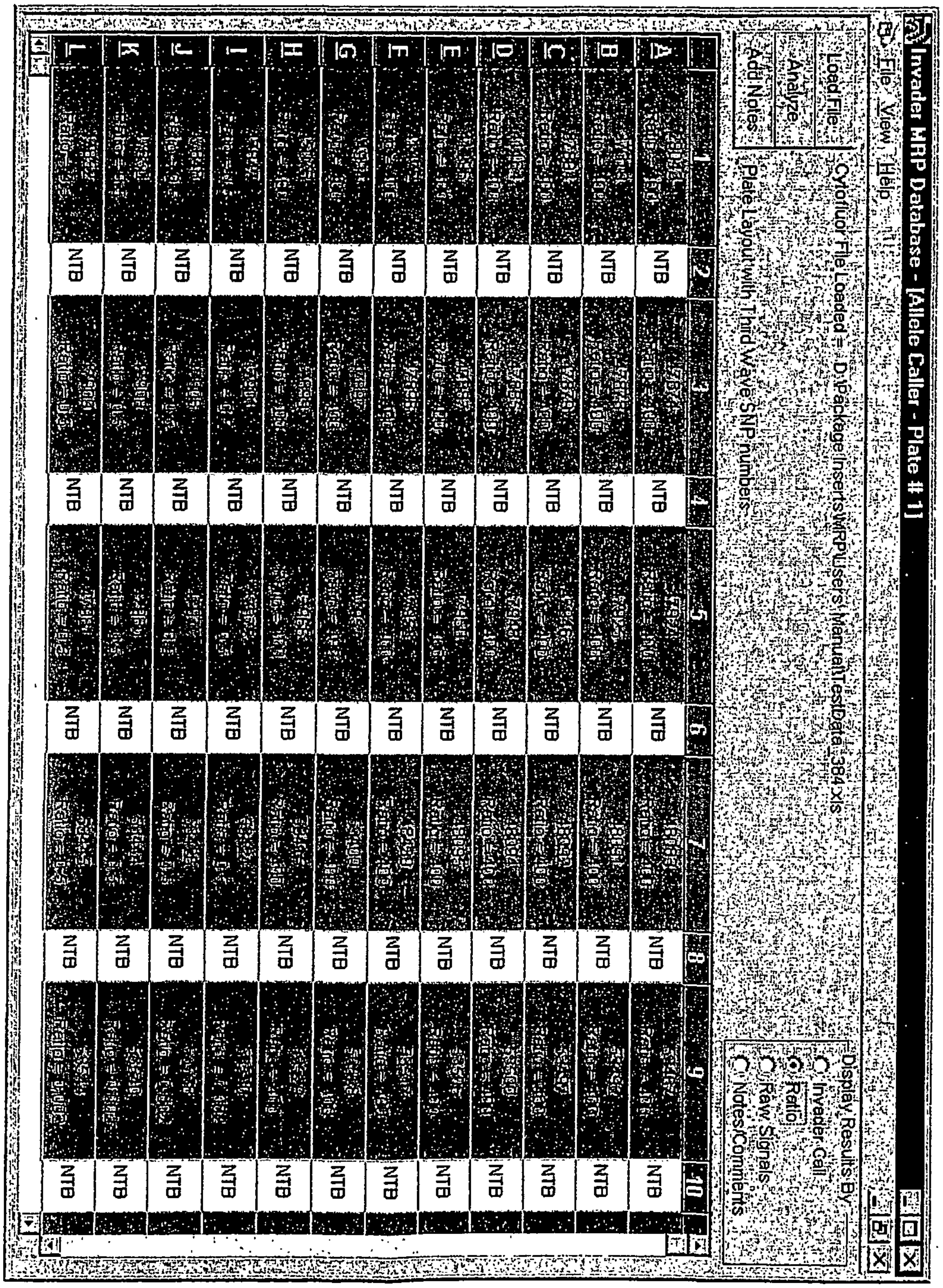

Invader MRP Database - [Allele Caller - Plate # 1]
File View Help
Load File
Analyze
Add Notes
Plate Layout with Third Wave SNP numbers
Display Results By
Invader Call
Ratio
Raw Signals
Notes/Comments
NTB

FIGURE 69

Microsoft Excel - 4_28Mar01.xls

A1 Snp number

	Snp number	Short Name	Fam Allele	Red Allele	Sample ID / Notes	Well	Raw Fam	Raw Red	Raw NTB Fam	Raw NTB Red	Fam FOZ	Red FOZ	FOZ RATIO	Dye Call	Final Call
2	29043	FMO1	G	A		A1	500	100	100	100	5	1	400	FAM	G
3	37319	ABCB2	C	T		B1	500	100	100	100	5	1	400	FAM	C
4	37804	MYH11	C	T		C1	500	100	100	100	5	1	400	FAM	C
5	41164	HGFAC	C	T		D1	500	100	100	100	5	1	400	FAM	C
6	41296	ASE-1	C	A		E1	500	100	100	100	5	1	400	FAM	C
7	41472	ECGF1	C	T		F1	500	100	100	100	5	1	400	FAM	C
8	44994	ALG6	G	A		G1	500	100	100	100	5	1	400	FAM	G
9	45076	GSTM4	C	G		H1	500	100	100	100	5	1	400	FAM	C
10	47847	ABCB1	G	A		I1	300	500	100	100	3	5	0.5	HET	HET
11	47848	ABCB1	G	A		J1	300	500	100	100	3	5	0.5	HET	HET
12	47850	ABCB1	C	T		K1	300	500	100	100	3	5	0.5	HET	HET
13	47852	ABCC8	C	T		L1	300	500	100	100	3	5	0.5	HET	HET
14	47854	ACADM	G	A		M1	300	500	100	100	3	5	0.5	HET	HET
15	47855	ADRB2	G	A		N1	300	500	100	100	3	5	0.5	HET	HET
16	47857	ADRB2	C	T		O1	300	500	100	100	3	5	0.5	HET	HET
17	47864	ALDH2	G	A		P1	300	500	100	100	3	5	0.5	HET	HET
18	47874	CASR	G	T		A3	500	100	100	100	5	1	400	FAM	G
19	47878	CETP	G	A		B3	500	100	100	100	5	1	400	FAM	G
20	47879	CETP	G	A		C3	500	100	100	100	5	1	400	FAM	G
21	47885	CYP2E	T	A		D3	500	100	100	100	5	1	400	FAM	T
22	47887	DIA4	C	T		E3	500	100	100	100	5	1	400	FAM	C
23	47896	FABP2	G	A		F3	500	100	100	100	5	1	400	FAM	G
24	47897	FMO3	G	A		G3	500	100	100	100	5	1	400	FAM	G
25	47899	GCG	G	A		H3	500	100	100	100	5	1	400	FAM	G
26	47901	GCGR	G	A		I3	300	500	100	100	3	5	0.5	HET	HET
27	47906	GYS1	G	A		J3	300	500	100	100	3	5	0.5	HET	HET
28	47908	HNF4A	C	T		K3	300	500	100	100	3	5	0.5	HET	HET
29	47909	HNF4A	C	T		L3	300	500	100	100	3	5	0.5	HET	HET
30	47913	INS	T	A		M3	300	500	100	100	3	5	0.5	HET	HET
31	47915	INSR	G	A		N3	300	500	100	100	3	5	0.5	HET	HET
32	47917	INSR	G	A		O3	300	500	100	100	3	5	0.5	HET	HET

Cytofluor File: 1061-42E
Date: Tue Feb 20 08:53:43 2001
Plate Name: MJ384, Read Pattern: Quadrant

IDAW / Rules / Data

Ready NUM

FIGURE 70

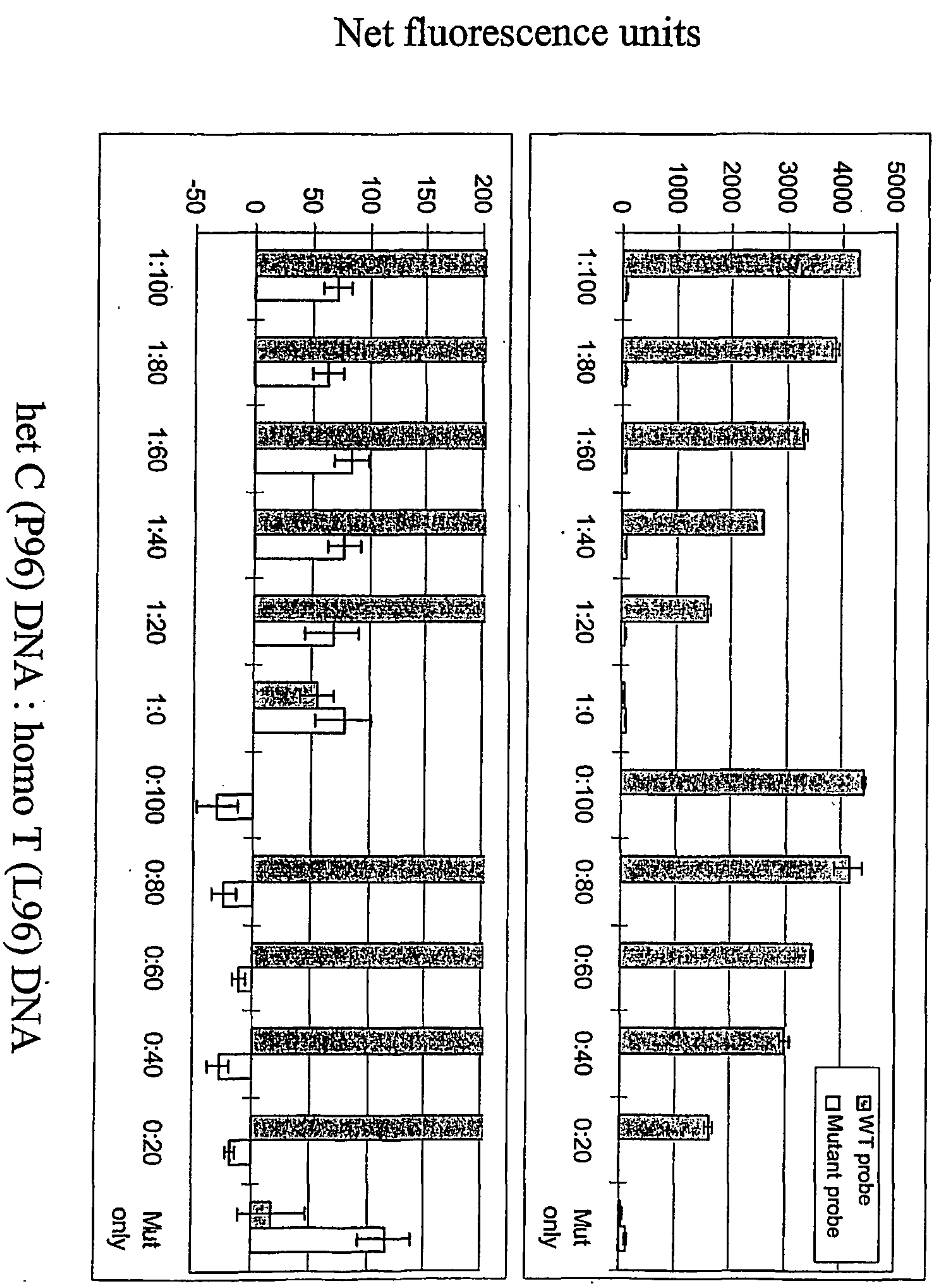

Net fluorescence units
WT probe
Mutant probe
5000
4000
3000
2000
1000
0
200
150
100
50
0
-50
1:100
1:80
1:60
1:40
1:20
1:0
0:100
0:80
0:60
0:40
0:20
Mut only
het C (P96) DNA : homo T (L96) DNA

FIGURE 71

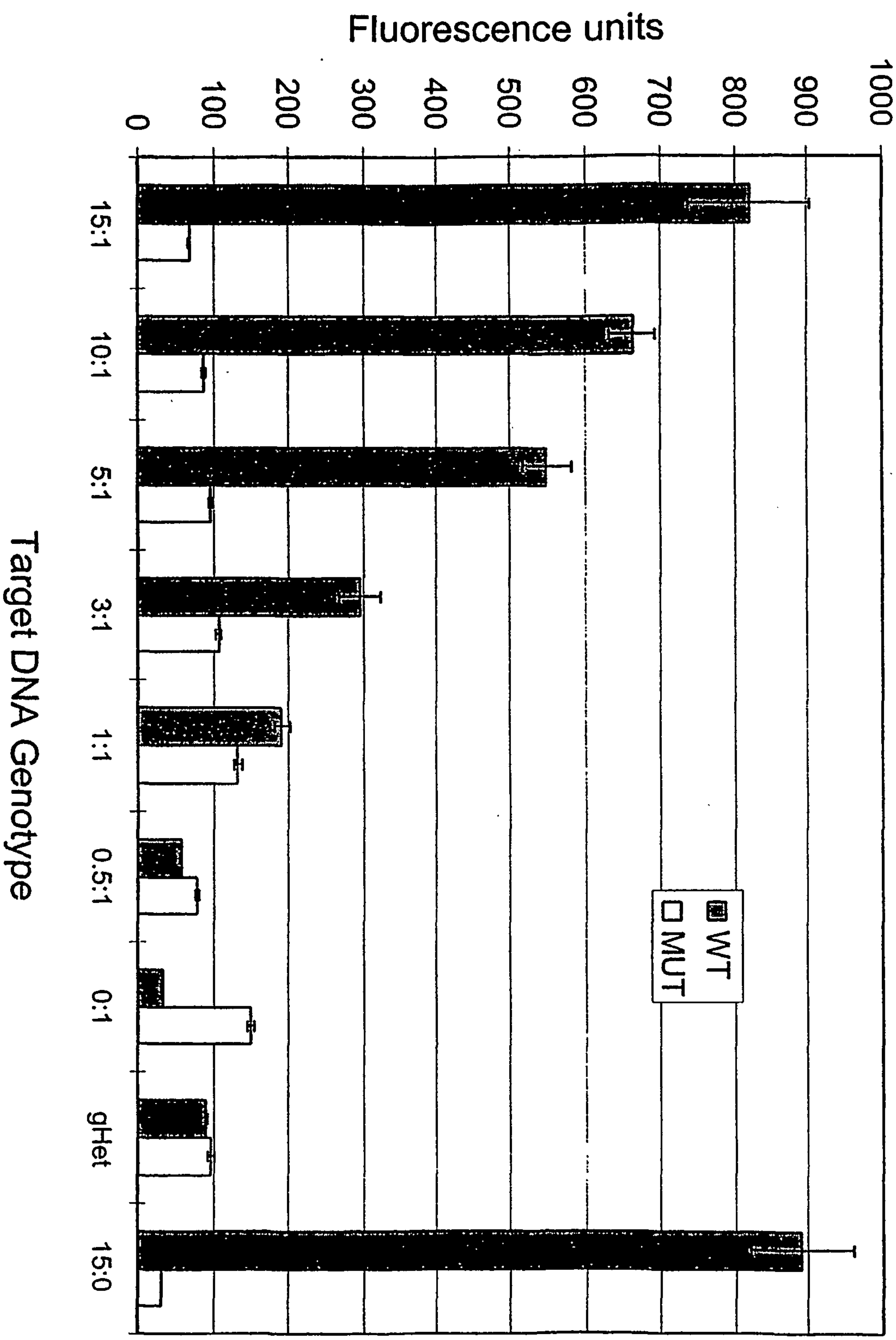

Fluorescence units
0
100
200
300
400
500
600
700
800
900
1000
WT
MUT
15:1
10:1
5:1
3:1
1:1
0.5:1
0:1
gHet
15:0
Target DNA Genotype

FIGURE 72

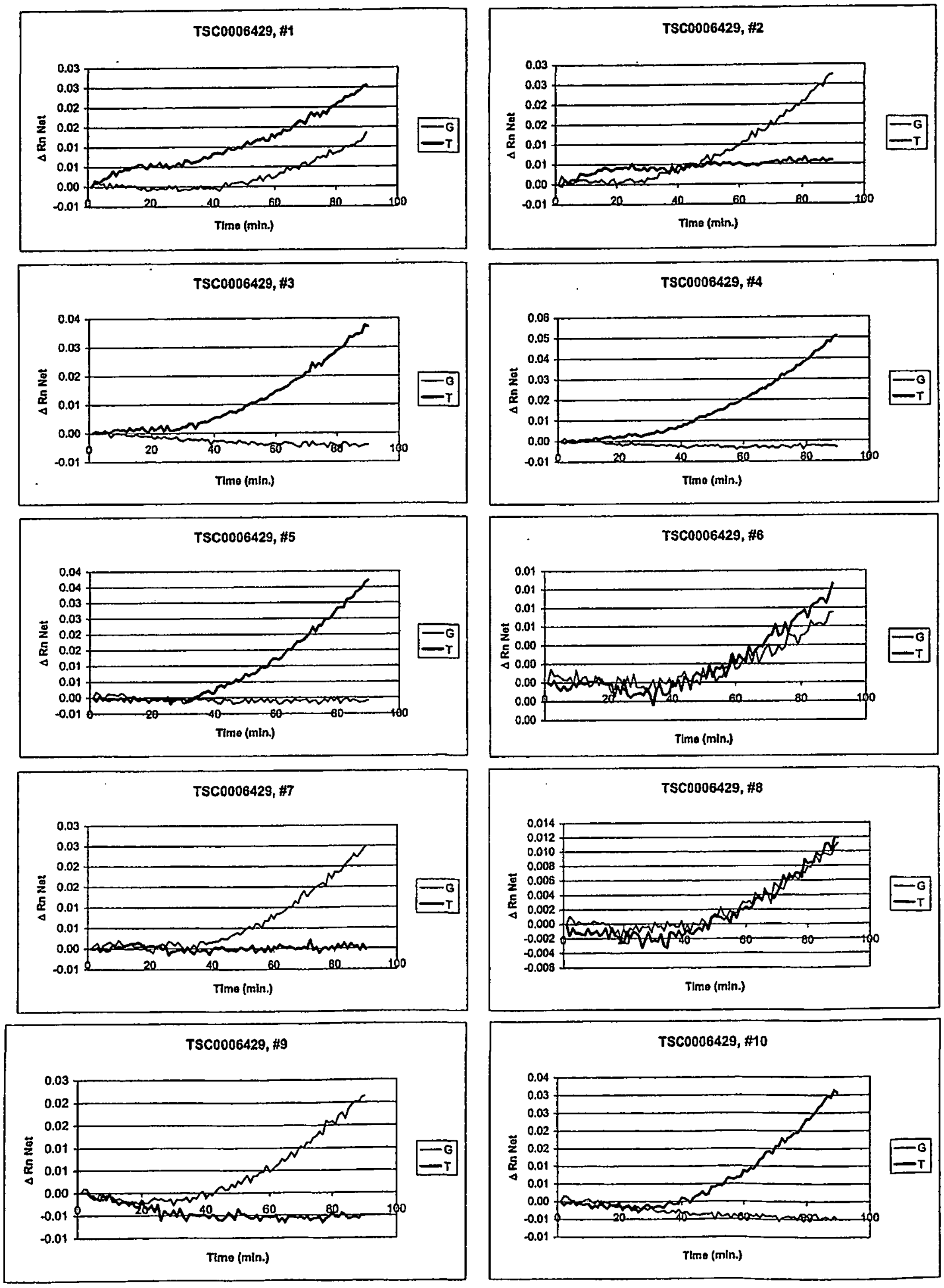

TSC0006429, #1
TSC0006429, #2
TSC0006429, #3
TSC0006429, #4
TSC0006429, #5
TSC0006429, #6
TSC0006429, #7
TSC0006429, #8
TSC0006429, #9
TSC0006429, #10
Δ Rn Net
Time (min.)
G
T

FIGURE 73

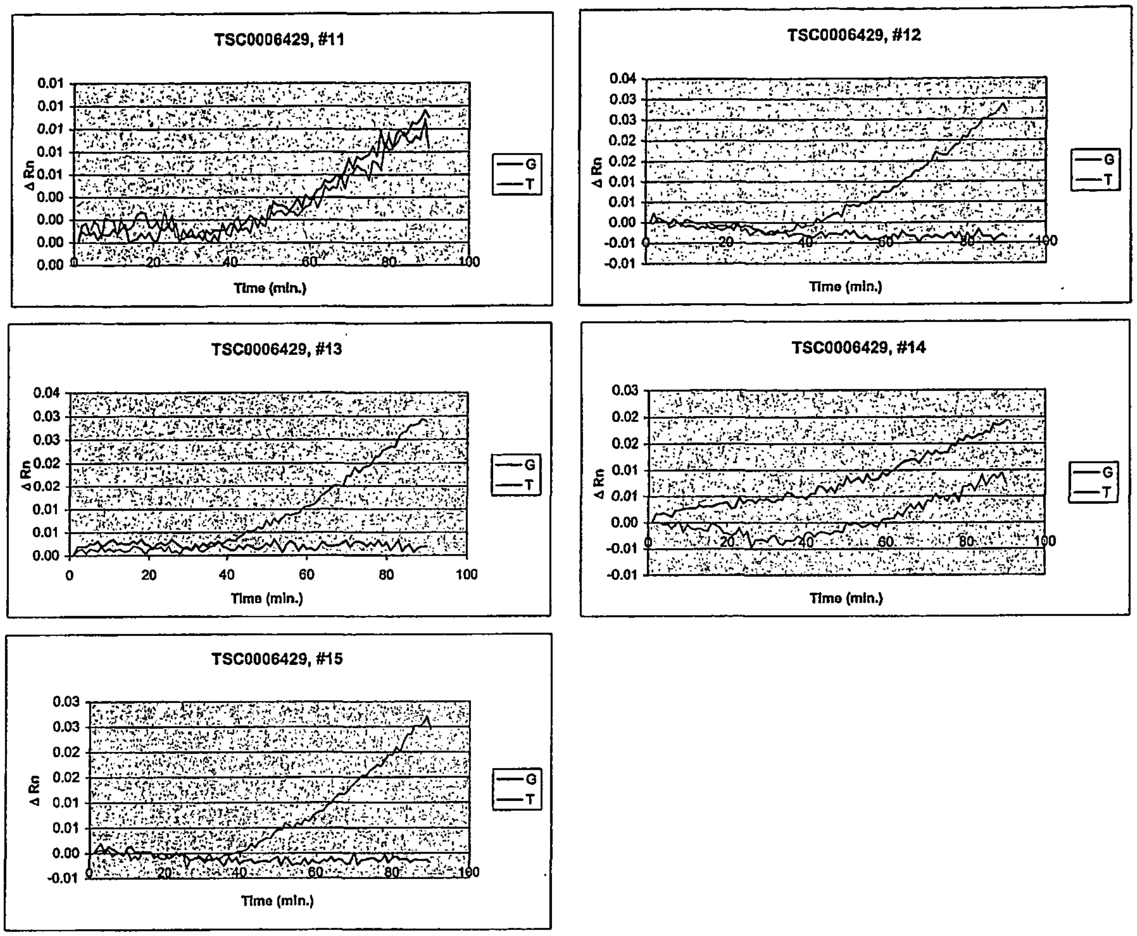

TSC0006429, #11
Δ Rn
Time (min.)
G
T
TSC0006429, #12
Δ Rn
Time (min.)
G
T
TSC0006429, #13
Δ Rn
Time (min.)
G
T
TSC0006429, #14
Δ Rn
Time (min.)
G
T
TSC0006429, #15
Δ Rn
Time (min.)
G
T

FIGURE 74

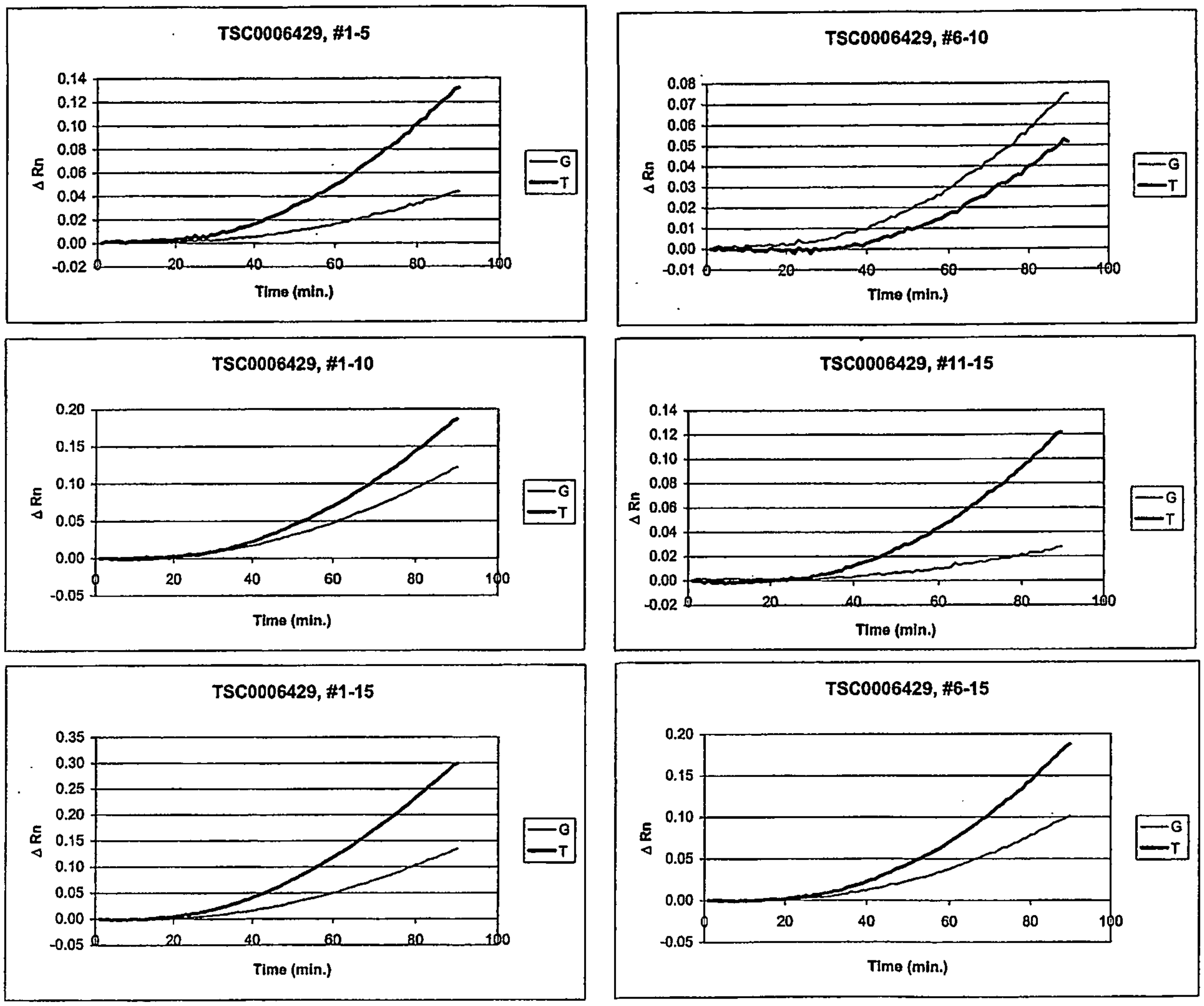
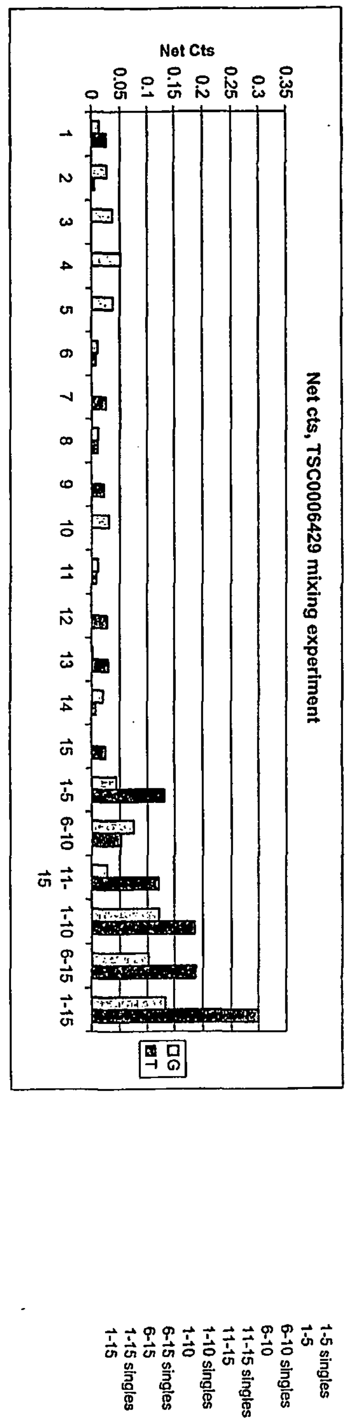
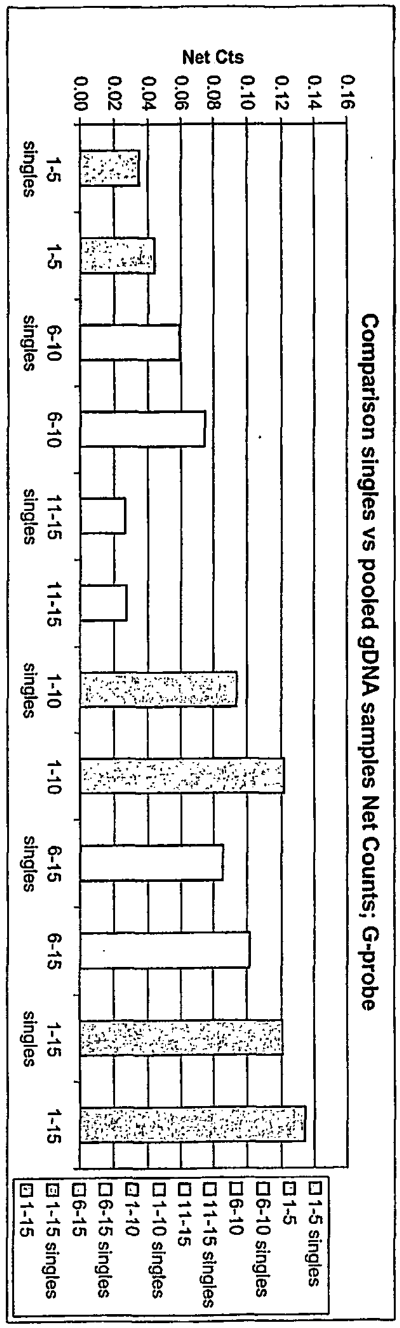

TSC0006429, #1-5
TSC0006429, #6-10
TSC0006429, #1-10
TSC0006429, #11-15
TSC0006429, #1-15
TSC0006429, #6-15
Δ Rn
Time (min.)
G
T

FIGURE 75


Net cts, TSC0006429 mixing experiment
Net Cts
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
1-5
6-10
11-15
1-10
6-15
1-15
G
T
1-5 singles
1-5
6-10 singles
6-10
11-15 singles
11-15
1-10 singles
1-10
6-15 singles
6-15
1-15 singles
1-15


Comparison singles vs pooled gDNA samples Net Counts; G-probe
Net Cts
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
1-5 singles
1-5
6-10 singles
6-10
11-15 singles
11-15
1-10 singles
1-10
6-15 singles
6-15
1-15 singles
1-15

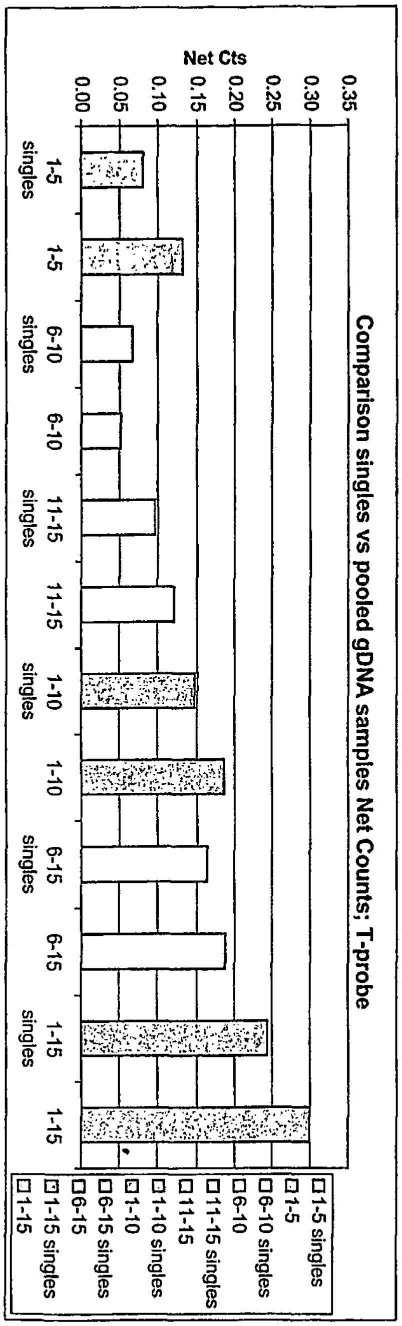

Comparison singles vs pooled gDNA samples Net Counts; T-probe
Net Cts
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
1-5 singles
1-5
6-10 singles
6-10
11-15 singles
11-15
1-10 singles
1-10
6-15 singles
6-15
1-15 singles
1-15

Figure 76

A

SNP #	0%	10%	20%	30%	50%	70%	80%	90%	100%
131809	1.2%	11.1%	20.2%	30.3%	49.2%	67.1%	79.0%	91.2%	97.0%
132140	-1.5%	12.2%	23.7%	34.8%	52.0%	69.7%	79.7%	89.2%	98.4%
132097	1.1%	10.8%	23.5%	33.7%	51.5%	69.0%	79.9%	88.1%	100.5%
132030	0.0%	13.1%	24.7%	34.4%	50.2%	66.2%	76.8%	87.6%	99.9%
131559	2.2%	12.4%	21.2%	37.1%	53.2%	72.1%	80.5%	90.8%	101.4%
131984	0.2%	13.0%	23.8%	34.2%	52.5%	70.8%	81.8%	90.8%	98.9%
132204	-1.9%	9.7%	21.0%	32.1%	52.1%	69.8%	80.5%	90.4%	102.3%
131534	-3.4%	9.2%	20.3%	30.9%	50.1%	68.8%	78.2%	89.2%	98.9%
131351	-1.0%	11.0%	21.9%	31.5%	49.5%	66.2%	75.1%	86.1%	97.6%
131532	-4.9%	8.0%	23.4%	35.0%	53.3%	70.4%	79.1%	87.2%	96.9%
132079	-5.5%	9.5%	21.3%	31.8%	49.2%	68.2%	77.9%	88.8%	100.4%
132505	-0.2%	10.5%	22.3%	31.6%	49.4%	66.8%	77.1%	87.1%	99.6%
132939	0.1%	12.1%	24.1%	33.4%	51.7%	70.0%	79.3%	89.6%	100.5%
Average	-1.0%	11.0%	22.4%	33.1%	51.1%	68.8%	78.8%	88.9%	99.4%
SD	2.3%	1.6%	1.5%	1.9%	1.5%	1.9%	1.8%	1.6%	1.7%

B

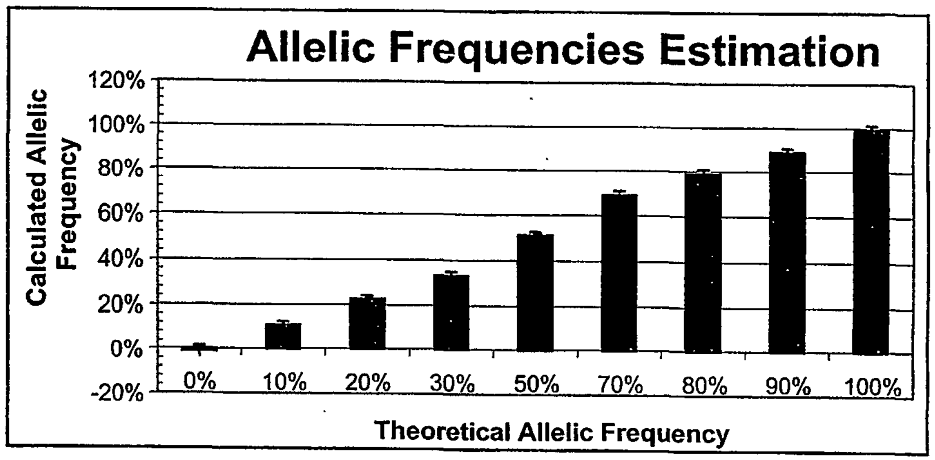
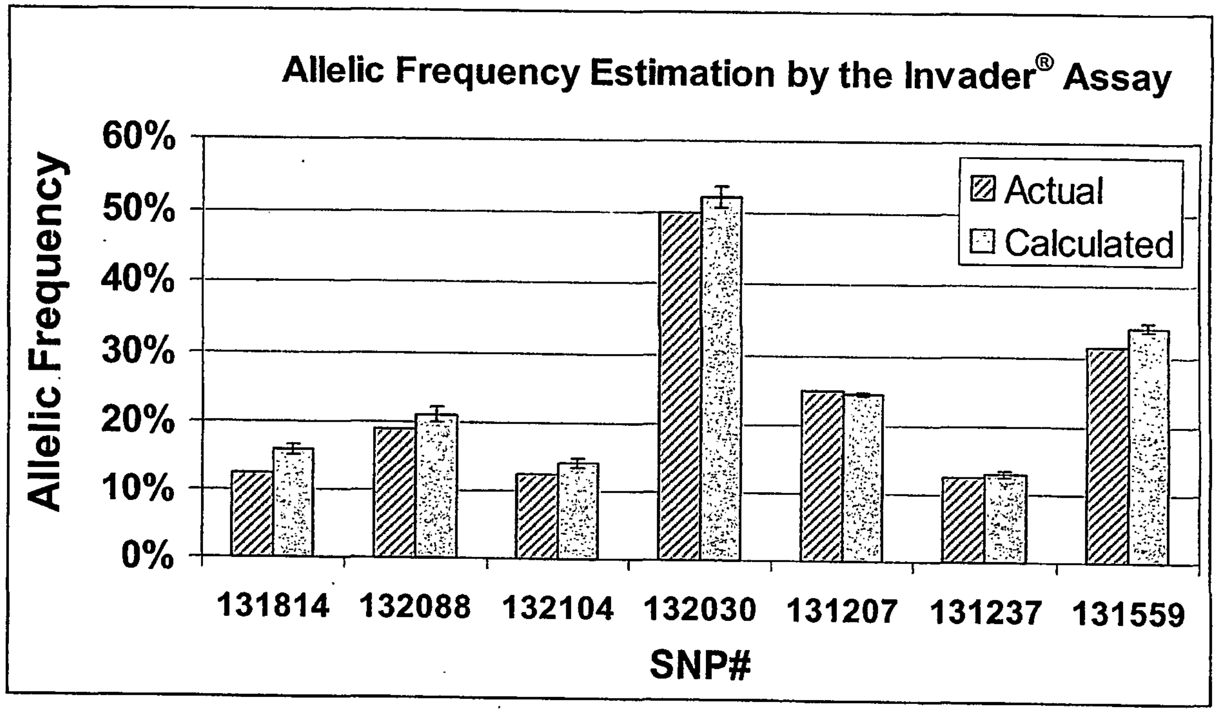
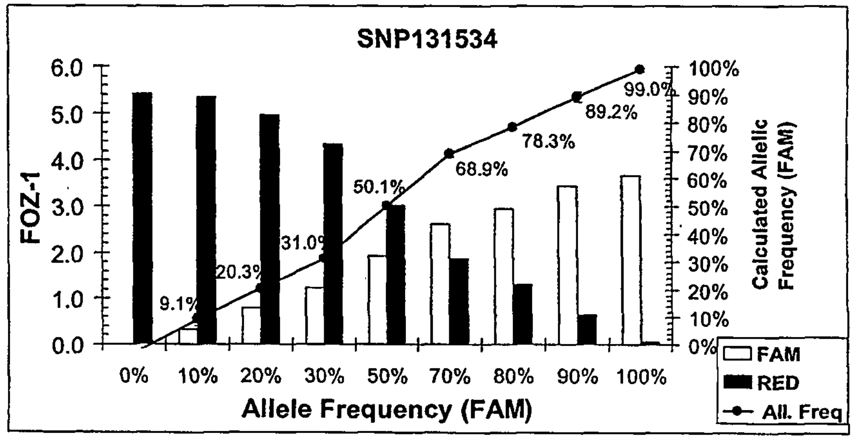


Figure 77



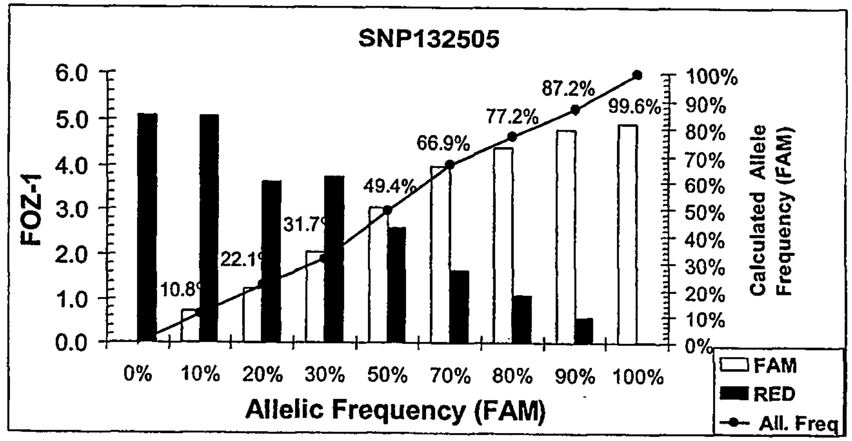
	131814	132088	132104	132030	131207	131237	131559
Actual	12.5%	18.8%	12.5%	50.0%	25.0%	12.5%	31.3%
Calculated	15.9%	21.0%	14.2%	52.2%	24.4%	13.0%	34.2%
SD	0.7%	1.1%	0.7%	1.5%	0.3%	0.5%	0.7%

Figure 78



	FAM Allele Frequency								
	0%	10%	20%	30%	50%	70%	80%	90%	100%
Rep.1	-3.8%	10.3%	20.4%	30.9%	50.5%	69.0%	79.1%	89.4%	100.7%
Rep.2	-3.8%	8.1%	19.4%	30.1%	49.3%	69.3%	78.4%	90.5%	100.0%
Rep.3	-4.3%	9.1%	19.5%	31.0%	49.5%	67.7%	78.0%	88.1%	98.0%
Rep.4	-1.7%	9.1%	22.2%	31.9%	51.1%	69.4%	77.6%	89.0%	97.5%
Avg	-3.4%	9.1%	20.3%	31.0%	50.1%	68.9%	78.3%	89.2%	99.0%
SD	1.1%	0.9%	1.3%	0.7%	0.8%	0.8%	0.6%	1.0%	1.5%

Figure 79



	FAM Allele Frequency								
	0%	10%	20%	30%	50%	70%	80%	90%	100%
Rep.1	-0.8%	12.0%	20.3%	30.8%	48.6%	66.4%	76.5%	86.1%	99.4%
Rep.2	-0.1%	11.3%	22.3%	31.7%	49.8%	67.7%	77.2%	88.7%	98.8%
Rep.3	-0.8%	7.6%	21.7%	31.6%	49.3%	65.8%	78.1%	85.4%	100.6%
Rep.4	0.8%	12.1%	24.2%	32.5%	50.0%	67.8%	76.9%	88.8%	99.6%
Avg	**-0.2%**	**10.8%**	**22.1%**	**31.7%**	**49.4%**	**66.9%**	**77.2%**	**87.2%**	**99.6%**
SD	**0.8%**	**2.2%**	**1.6%**	**0.7%**	**0.6%**	**1.0%**	**0.7%**	**1.7%**	**0.8%**

Figure 80A

SNP#	Dye	OligoType	Poly	Oligonucleotide sequence	
131551		INVADER		CAAGCTCCACACTGGAAGAATCAGGAGCAATAGATTTCTTT	SEQ ID NO:18
	Red	Probe	G	CGCGCCGAGGCTTGCTCTCAGAAGGAAACY	SEQ ID NO:19
	FAM	Probe	C	ATGACGTGGCAGACGTTGCTCTCAGAAGGAAACY	SEQ ID NO:20
	Red	Target	G	GGTGGTTTCCTTCTGAGAGCAAGAAGAAATCTATTGCTCCTGATTCTTCCAGTGTGGAGCTTGGA	SEQ ID NO:21
	FAM	Target	C	GGTGGTTTCCTTCTGAGAGCAACAAGAAATCTATTGCTCCTGATTCTTCCAGTGTGGAGCTTGGA	SEQ ID NO:22
131532		INVADER		GCAATGAAATTGCACATTTTACCCAGCCATATGCCATGT	SEQ ID NO:23
	FAM	Probe	C	ATGACGTGGCAGACGGCAGCCAACATCAGY	SEQ ID NO:24
	Red	Probe	T	CGCGCCGAGGAGCAGCCAACATCAGTY	SEQ ID NO:25
	FAM	Target	C	AAACACTGATGTTGGCTGCCCATGGCATATGGCTGGGTAAAATGTGCAATTTCATTGCTT	SEQ ID NO:26
	Red	Target	T	AAACACTGATGTTGGCTGCTCATGGCATATGGCTGGGTAAAATGTGCAATTTCATTGCTT	SEQ ID NO:27
131534		INVADER		GGAGAGATACTAGGCACTCACTTCATACAAAAGAAAAACCAATGCTT	SEQ ID NO:28
	Red	Probe	A	CGCGCCGAGGATAGGCCTTATAAATGACTCTCY	SEQ ID NO:29
	FAM	Probe	G	ATGACGTGGCAGACGTAGGCCTTATAAATGACTCTCY	SEQ ID NO:30
	Red	Target	A	CTTTGAGAGTCATTTATAAGGCCTATAGCATTGGTTTTTCTTTTGTATGAAGTGAGTGCCTAGTATCTCTCCAC	SEQ ID NO:31
	FAM	Target	G	CTTTGAGAGTCATTTATAAGGCCTACAGCATTGGTTTTTCTTTTGTATGAAGTGAGTGCCTAGTATCTCTCCAC	SEQ ID NO:32
131809		INVADER		AGCACTTTACAATGGTTCAGCTTCCAATTTATATGCCAGATATTACTAAATACAGAGT	SEQ ID NO:33
	FAM	Probe	C	ATGACGTGGCAGACGTACCGTACTCTATAAAGTAAATGY	SEQ ID NO:34
	Red	Probe	T	CGCGCCGAGGATACCGTACTCTATAAAGTAAATGCY	SEQ ID NO:35
	FAM	Target	C	AGTTGCATTTACTTTATAGAGTACGGTACCTCTGTATTTAGTAATATCTGGCATATAAATTGGAAGCTGAACCATTGTAAAGTGCTAA	SEQ ID NO:36
	Red	Target	T	AGTTGCATTTACTTTATAGAGTACGGTATCTCTGTATTTAGTAATATCTGGCATATAAATTGGAAGCTGAACCATTGTAAAGTGCTAA	SEQ ID NO:37
131559		INVADER		CCCTCATCCTGGTTCAGAAATAACCGCGTGGT	SEQ ID NO:38
	Red	Probe	C	CGCGCCGAGGCATTGCTGTTGTTGCCY	SEQ ID NO:39
	FAM	Probe	G	ATGACGTGGCAGACGATTGCTGTTGTTGCCY	SEQ ID NO:40
	Red	Target	C	CAACGGCAACAACAGCAATGCCACGCGGTTATTTCTGAACCAGGATGAGGGTG	SEQ ID NO:41
	FAM	Target	G	CAACGGCAACAACAGCAATCCCACGCGGTTATTTCTGAACCAGGATGAGGGTG	SEQ ID NO:42
131984		INVADER		GTCCTCAATGATGGGAGGGCATTCACCTCTACAT	SEQ ID NO:43
	Red	Probe	T	CGCGCCGAGGAGAAAGAAGACAGATGGTCAY	SEQ ID NO:44
	FAM	Probe	C	ATGACGTGGCAGACGGAAAGAAGACAGATGGTCY	SEQ ID NO:45
	Red	Target	T	AGGTTGACCATCTGTCTTCTTTCTTGTAGAGGTGAATGCCCTCCCATCATTGAGGACAG	SEQ ID NO:46
	FAM	Target	C	AGGTTGACCATCTGTCTTCTTTCCTGTAGAGGTGAATGCCCTCCCATCATTGAGGACAG	SEQ ID NO:47

Figure 80B

SNP#	Dye	OligoType	Poly	Oligonucleotide sequence
132079		INVADER		GCCAGGCTTGTAAATTACATGAGCAGCCCTCTT
	FAM	Probe	G	ATGACGTGGCAGACGCAAAGGGAGTCTAAACCY
	Red	Probe	C	CGCGCCGAGGCCAAAGGGAGTCTAAACCY
	FAM	Target	G	TTCAGGTTTAGACTCCCTTTGCAGAGGGCTGCTCATGTAATTTACAAGCCTGGCAG
	Red	Target	C	TTCAGGTTTAGACTCCCTTTGGAGAGGGCTGCTCATGTAATTTACAAGCCTGGCAG
132097		INVADER		GTGACACAGGCCAGTAAGTTACTCAAATTTTAAGTTTGAGCTTTTTCAA
	Red	Probe	T	CGCGCCGAGGTTTGTAAGATGGAACAATCGTY
	FAM	Probe	C	ATGACGTGGCAGACCTTGTAAGATGGAACAATCGTY
	Red	Target	T	TGTCACGATTGTTCCATCTTACAAATGAAAAAGCTCAAACTTAAAATTTGAGTAACTTACTGGCCTGTGTCACAT
	FAM	Target	C	TGTCACGATTGTTCCATCTTACAAGTGAAAAAGCTCAAACTTAAAATTTGAGTAACTTACTGGCCTGTGTCACAT
132204		INVADER		GTGTTACAAGCTGCCTCTCCAAAATCAATGCCTTCACTATAT
	FAM	Probe	G	ATGACGTGGCAGACGAACTTGCCTGAAGCAAY
	Red	Probe	C	CGCGCCGAGGCAACTTGCCTGAAGCAY
	FAM	Target	G	GTCATTGCTTCAGGCAAGTTCTATAGTGAAGGCATTGATTTTGGAGAGGCAGCTTGTAACACGT
	Red	Target	C	GTCATTGCTTCAGGCAAGTTGTATAGTGAAGGCATTGATTTTGGAGAGGCAGCTTGTAACACGT
132030		INVADER		GCCTTTTCAAATCGACTTTCTCAAATGTTTTGCCTGTTCTCTCT
	Red	Probe	T	CGCGCCGAGGATTTCCATTCCCAGTGCY
	FAM	Probe	C	ATGACGTGGCAGACGTTTCCATTCCCAGTGCY
	Red	Target	T	TAATGCACTGGGAATGGAAATGAGAGAACAGGCAAAACATTTGAGAAAGTCGATTTGAAAAGGCAG
	FAM	Target	C	TAATGCACTGGGAATGGAAACGAGAGAACAGGCAAAACATTTGAGAAAGTCGATTTGAAAAGGCAG
132140		INVADER		GGTATTGTGTGTCCAGTTTTGTTTGTAAAATGTAACCTTCGTGTGAATGT
	Red	Probe	T	CGCGCCGAGGACCATCTATTCCTTTCTTTTTGY
	FAM	Probe	C	ATGACGTGGCAGACGCCATCTATTCCTTTCTTTTTGY
	Red	Target	T	CTTCCAAAAAGAAAGGAATAGATGGTCATTCACACGAAGGTTACATTTTACAAACAAAACTGGACACACAATACCAT
	FAM	Target	C	CTTCCAAAAAGAAAGGAATAGATGGCCATTCACACGAAGGTTACATTTTACAAACAAAACTGGACACACAATACCAT
132505		INVADER		ACCACCTGCCTCTCATCCATGCAGCAAC
	Red	Probe	T	CGCGCCGAGGTTCATCAGGCCTGTATATAAAAY
	FAM	Probe	A	ATGACGTGGCAGACATCATCAGGCCTGTATATAAAAY
	Red	Target	T	CTTGTTTTATATACAGGCCTGATGAATTGCTGCATGGATGAGAGGCAGGTGGTGG
	FAM	Target	A	CTTGTTTTATATACAGGCCTGATGATTTGCTGCATGGATGAGAGGCAGGTGGTGG

Figure 80C

SNP#	Dye	OligoType	Poly	Oligonucleotide sequence	
132939		INVADER		GCCAAGCAGTGGTCATGAAAGTCCAGCCT	SEQ ID NO:78
	FAM	Probe	G	ATGACGTGGCAGACGCTGTCACATCCTTGAGY	SEQ ID NO:79
	Red	Probe	C	CGCGCCGAGGCCTGTCACATCCTTGAGY	SEQ ID NO:80
	FAM	Target	G	GTTCCTCAAGGATGTGACAGCGGCTGGACTTTCATGACCACTGCTTGGCCA	SEQ ID NO:81
	Red	Target	C	GTTCCTCAAGGATGTGACAGGGGCTGGACTTTCATGACCACTGCTTGGCCA	SEQ ID NO:82
131814		INVADER		CTCTGTGCCATCTTTTACTCCATGAACTGCATTTAATGTGTAGCT	SEQ ID NO:83
	Red	Probe	T	CGCGCCGAGGATCTGCATTTTTCAACTGGTY	SEQ ID NO:84
	FAM	Probe	C	ATGACGTGGCAGACGTCTGCATTTTTCAACTGGY	SEQ ID NO:85
	Red	Target	T	AGACACCAGTTGAAAAATGCAGATGCTACACATTAAATGCAGTTCATGGAGTAAAAGATGGCACAGAGCT	SEQ ID NO:86
	FAM	Target	C	AGACACCAGTTGAAAAATGCAGACGCTACACATTAAATGCAGTTCATGGAGTAAAAGATGGCACAGAGCT	SEQ ID NO:87
131207		INVADER		TGCGCAAACTGGTTTAATATCATTAGTGTAACAGCCAAGGTGT	SEQ ID NO:88
	Red	Probe	A	CGCGCCGAGGATCAAAGGCAGTAAATTATAAACTTY	SEQ ID NO:89
	FAM	Probe	C	ATGACGTGGCAGACCTCAAAGGCAGTAAATTATAAACTY	SEQ ID NO:90
	Red	Target	A	CTACAAGTTTATAATTTACTGCCTTTGATCACCTTGGCTGTTACACTAATGATATTAAACCAGTTTGCGCAAT	SEQ ID NO:91
	FAM	Target	C	CTACAAGTTTATAATTTACTGCCTTTGAGCACCTTGGCTGTTACACTAATGATATTAAACCAGTTTGCGCAAT	SEQ ID NO:92
131237		INVADER		AGCATCAAGCCTGCTACTAAAAATATTTTTTCCTGCTGCTCTGT	SEQ ID NO:93
	Red	Probe	A	CGCGCCGAGGAGAATGTGTGTTCTTCCATCY	SEQ ID NO:94
	FAM	Probe	G	ATGACGTGGCAGACGGAATGTGTGTTCTTCCATY	SEQ ID NO:95
	Red	Target	A	TTTAGATGGAAGAACACACATTCTCAGAGCAGCAGGAAAAAATATTTTTAGTAGCAGGCTTGATGCTTA	SEQ ID NO:96
	FAM	Target	G	TTTAGATGGAAGAACACACATTCCCAGAGCAGCAGGAAAAAATATTTTTAGTAGCAGGCTTGATGCTTA	SEQ ID NO:97
132030		INVADER		GCCTTTTCAAATCGACTTTCTCAAATGTTTTGCCTGTTCTCTCT	SEQ ID NO:98
	Red	Probe	T	CGCGCCGAGGATTTCCATTCCCAGTGCY	SEQ ID NO:99
	FAM	Probe	C	ATGACGTGGCAGACGTTTCCATTCCCAGTGCY	SEQ ID NO:100
	Red	Target	T	TAATGCACTGGGAATGGAAATGAGAGAACAGGCAAAACATTTGAGAAAGTCGATTTGAAAAGGCAG	SEQ ID NO:101
	FAM	Target	C	TAATGCACTGGGAATGGAAACGAGAGAACAGGCAAAACATTTGAGAAAGTCGATTTGAAAAGGCAG	SEQ ID NO:102
132104		INVADER		TTGAGTAACAAAGATTGGGCCCTTATACCTGTGAA	SEQ ID NO:103
	FAM	Probe	C	ATGACGTGGCAGACCGTGTCTGGCAATGACY	SEQ ID NO:104
	Red	Probe	T	CGCGCCGAGGTGTGTCTGGCAATGACY	SEQ ID NO:105
	FAM	Target	C	CTTTGTCATTGCCAGACACGTCACAGGTATAAGGGCCCAATCTTTGTTACTCAAAT	SEQ ID NO:106
	Red	Target	T	CTTTGTCATTGCCAGACACATCACAGGTATAAGGGCCCAATCTTTGTTACTCAAAT	SEQ ID NO:107
132088		INVADER		CAAATGGCATTTCAAATGCATAAAAATAACTTATTCGTAAATTTTCTTTCTCTCA	SEQ ID NO:108
	Red	Probe	T	CGCGCCGAGGTTTCTTGTTTTAGTATAGCACCTY	SEQ ID NO:109
	FAM	Probe	C	ATGACGTGGCAGACCTTCTTGTTTTAGTATAGCACCTY	SEQ ID NO:110
	Red	Target	T	TTCAAGGTGCTATACTAAAACAAGAAAGAGAGAAAGAAAATTTACGAATAAGTTATTTTTATGCATTTGAAATGCCATTTGGA	SEQ ID NO:111
	FAM	Target	C	TTCAAGGTGCTATACTAAAACAAGAAGGAGAGAAAGAAAATTTACGAATAAGTTATTTTTATGCATTTGAAATGCCATTTGGA	SEQ ID NO:112

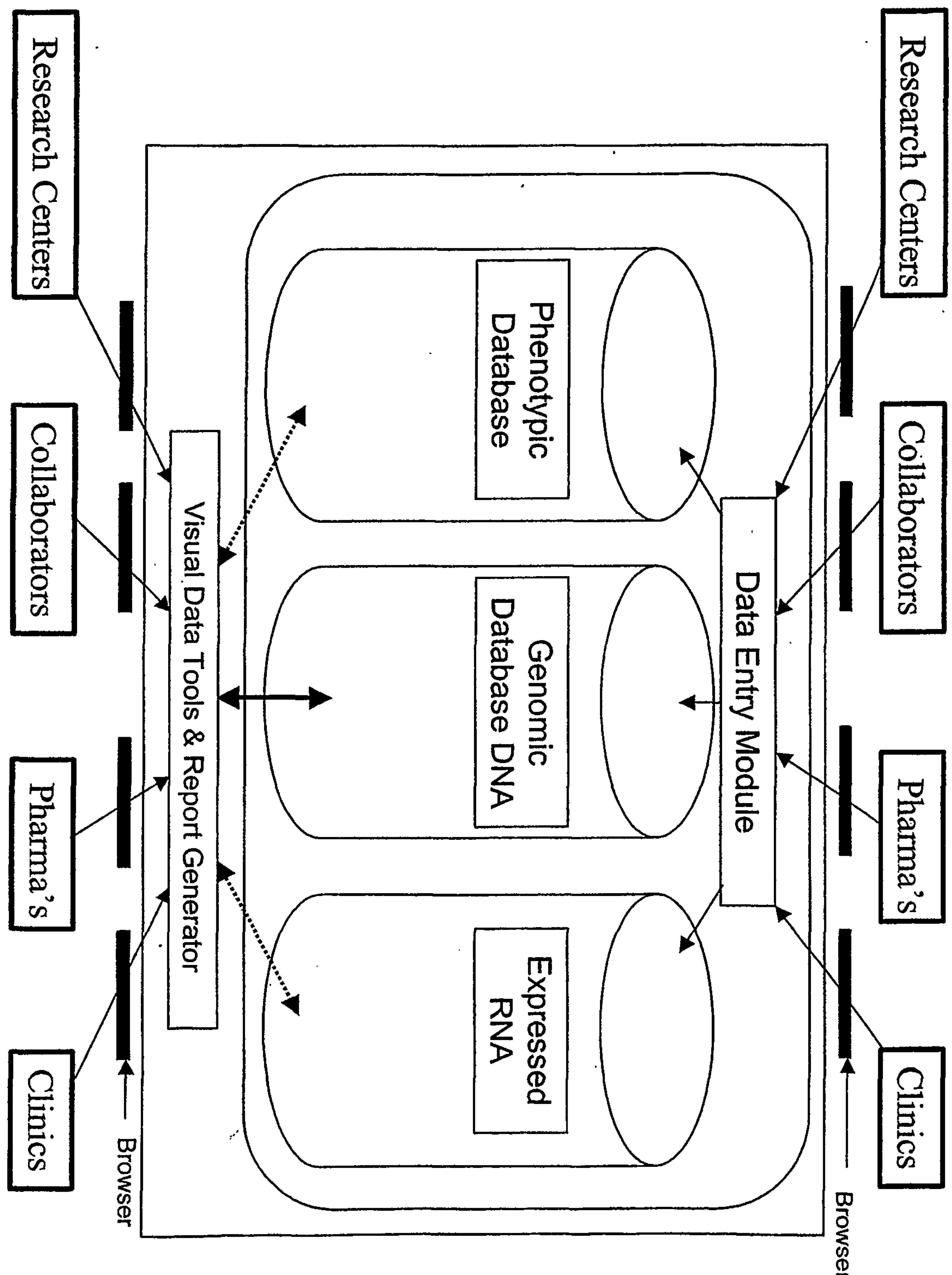

Research Centers
Collaborators
Pharma's
Clinics
Browser
Data Entry Module
Phenotypic Database
Genomic Database DNA
Expressed RNA
Visual Data Tools & Report Generator
Research Centers
Collaborators
Pharma's
Clinics
Browser

Figure 81

FIGURE 82

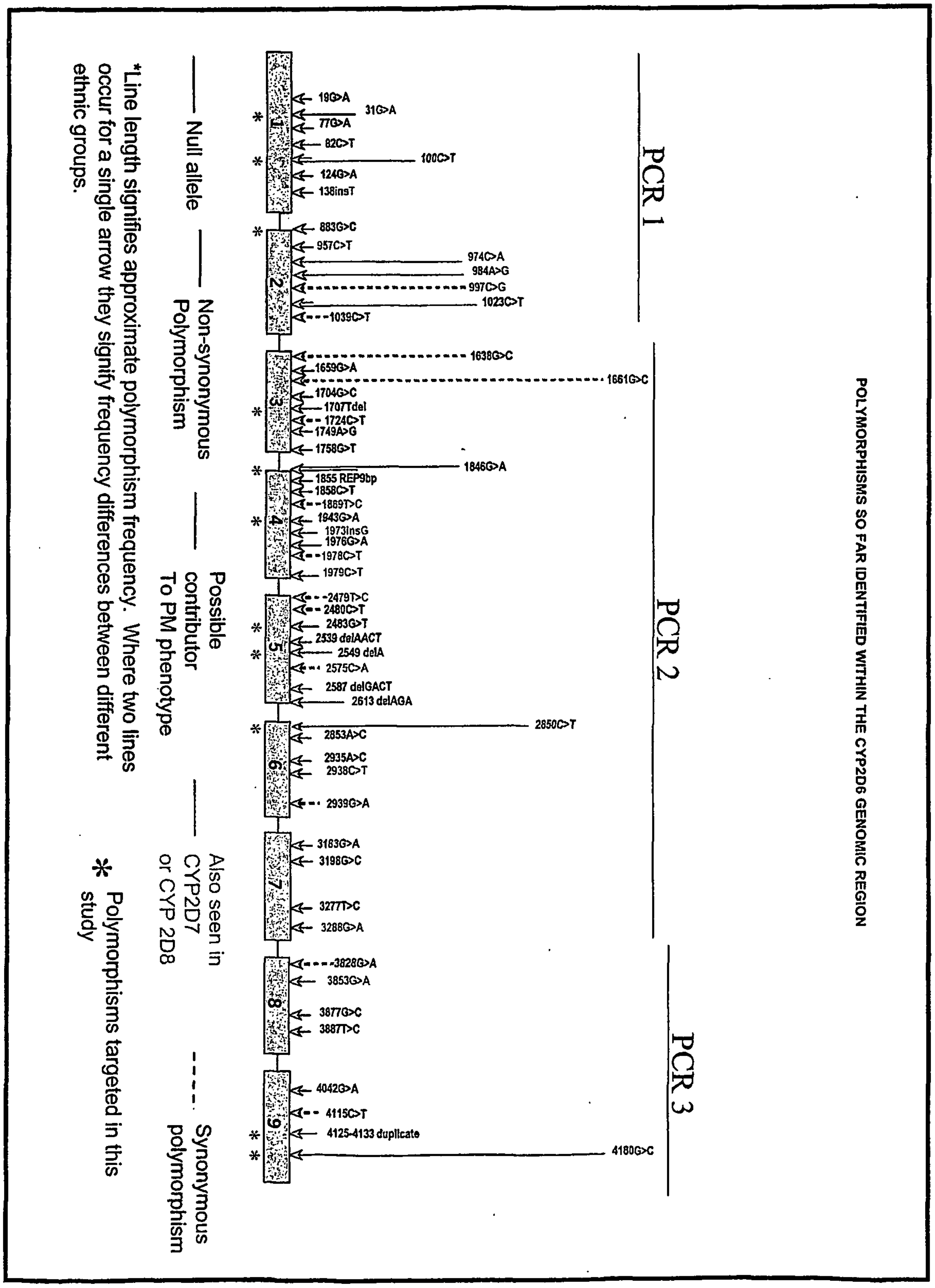

POLYMORPHISMS SO FAR IDENTIFIED WITHIN THE CYP2D6 GENOMIC REGION
PCR 1
PCR 2
PCR 3
1
2
3
4
5
6
7
8
9
19G>A
31G>A
77G>A
82C>T
100C>T
124G>A
138insT
883G>C
957C>T
974C>A
984A>G
997C>G
1023C>T
1039C>T
1638G>C
1659G>A
1661G>C
1704G>C
1707Tdel
1724C>T
1749A>G
1758G>T
1846G>A
1855 REP9bp
1858C>T
1869T>C
1943G>A
1973insG
1976G>A
1978C>T
1979C>T
2479T>C
2480C>T
2483G>T
2539 delAACT
2549 delA
2575C>A
2587 delGACT
2613 delAGA
2850C>T
2853A>C
2935A>C
2938C>T
2939G>A
3183G>A
3198G>C
3277T>C
3288G>A
3828G>A
3853G>A
3877G>C
3887T>C
4042G>A
4115C>T
4125-4133 duplicate
4180G>C
Null allele
Non-synonymous Polymorphism
Possible contributor To PM phenotype
Also seen in CYP2D7 or CYP 2D8
Synonymous polymorphism
*Line length signifies approximate polymorphism frequency. Where two lines occur for a single arrow they signify frequency differences between different ethnic groups.
* Polymorphisms targeted in this study

FIGURE 83

A - CYP2D6 SPECIFIC PCR AMPLIFICATION

The two CYP2D pseudogenes, CYP2D7 and CYP2D8 share many of the identified polymorphisms, and over 80% sequence similarity with the target gene CYP2D6. Therefore, to prevent false positive results, due to detection of the two pseudogenes, a CYP2D6 specific Triplex PCR amplification reaction has been developed to integrate with the Invader® assay. The three PCR products are amplified from genomic template in a single tube using CYP2D6 specific PCR primers with a 35 cycle PCR reaction of 95°C for 20 sec and 68°C 2 for min.

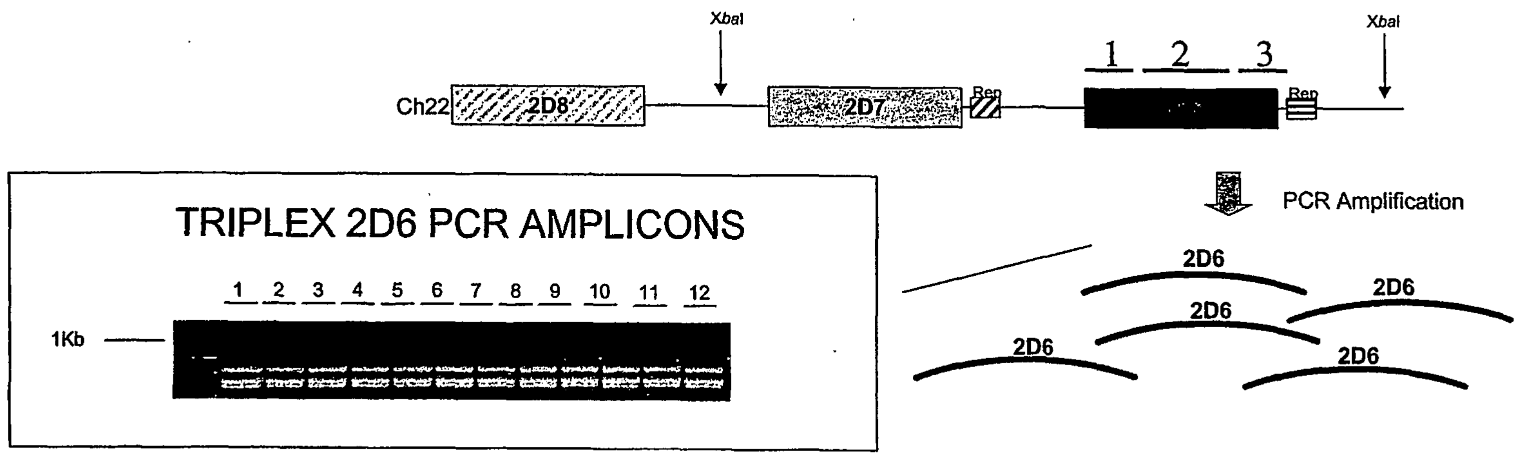


FIGURE 84

B - Biplex Signal detection

A 1/20 dilution of the CYP2D6 specific PCR products are used as a template for polymorphism detection using the Biplex Invader® assay system in a single well of a 96 or 384 well plate. Two serial Invader® reactions occur simultaneously, target detection / allele discrimination takes place in the primary Invader® reaction, whilst signal amplification takes place in the secondary Invader® reaction using a set of universal signal probes. The entire assay is isothermal and only requires a single step to set up. In addition to this, signal can be read and alleles called after only 20 minute incubation at 63°C following an initial 5 minute 95°C denaturation step.

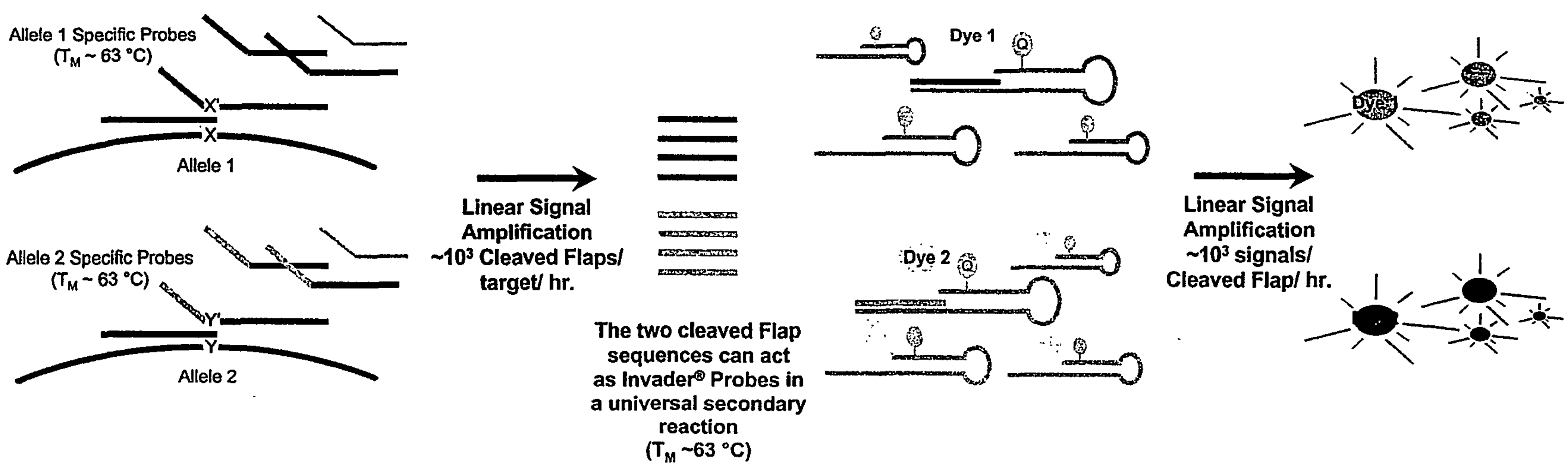


FIGURE 85

A Screen of 175 Individuals with 11 CYP2D6 INVADER Assays

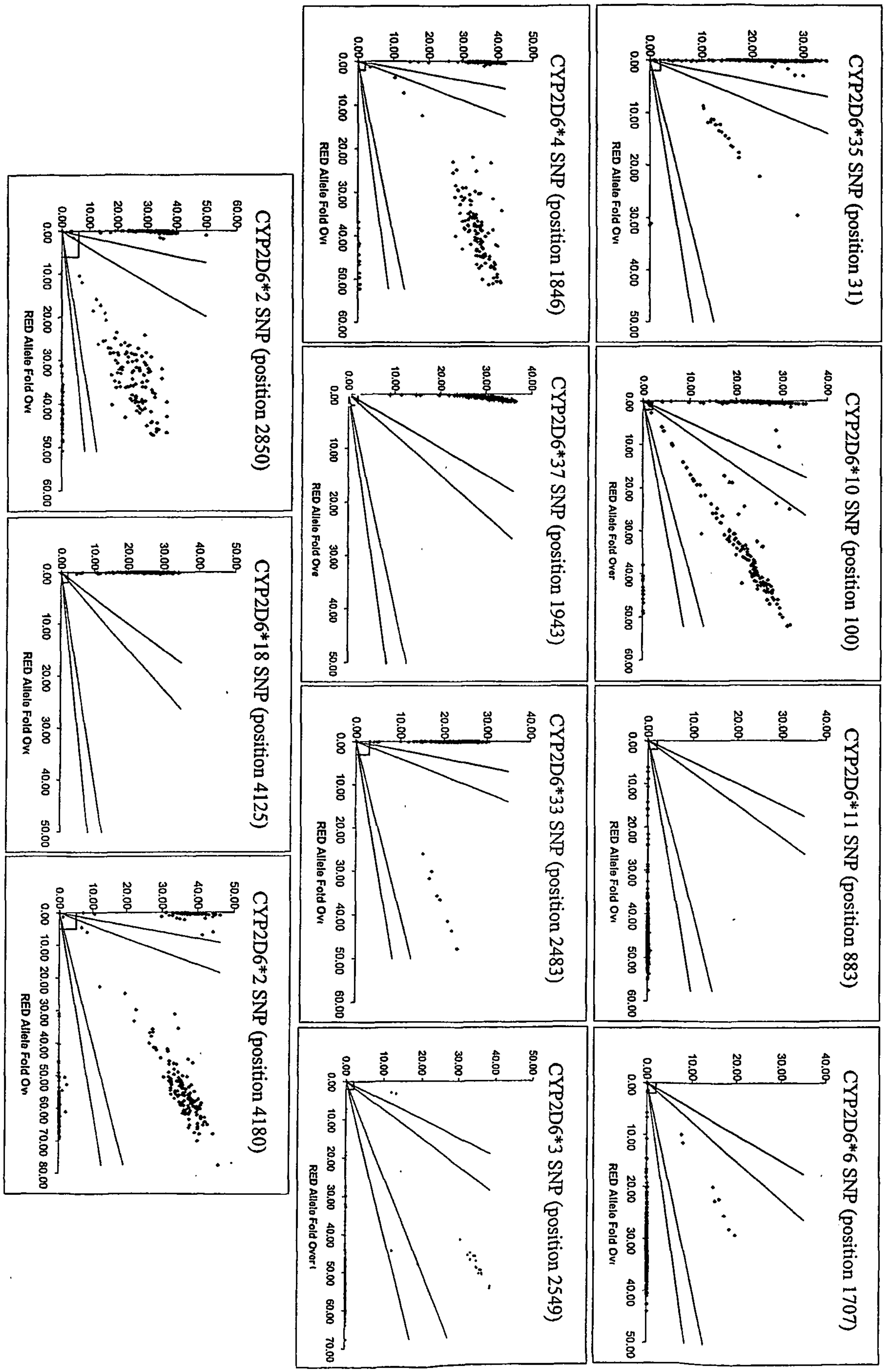


FIGURE 86

A Screen of 175 Individuals with CYP2D6 INVADER Assays

Position	31	100	883	1707	1846	1943	2483	2549	2850	4125-4133	4180
Allele	*35	*10	*11	*6	*4	*37	*33	*3	*2	*18	*2
PCR product	PCR1	PCR1	PCR1	PCR2	PCR2	PCR2	PCR2	PCR2	PCR2	PCR3	PCR3
Polymorphism	G>A	C>T	G>C	T>Del	G>A	G>A	G>T	A>Del	C>T	duplication	G>C
Effect	V11M	P34S	splice defect	frameshift	splice defect	R201H	A237S	frameshift	R296C	468VPT470	S486T
n=	174	173	175	175	174	175	174	174	173	175	174
WT frequency	159	99	175	171	103	175	170	165	88	175	41
HET frequency	14	66	0	4	64	0	4	8	61	0	77
Mut frequency	1	8	0	0	7	0	0	1	24	0	56
Hardy Weinberg	$\chi_2$1.2 (p=~0.15)	$\chi_2$0.5 (p=>0.2)		$\chi_2$0.0 (p=>0.2)	$\chi_2$0.6 (p=>0.2)		$\chi_2$0.0 (p=>0.2)	$\chi_2$5.4 (p=0.02)	$\chi_2$5.8 (p=~0.02)		$\chi_2$2 (p=~0.15)
Rare allele frequency	4.79	29.3	0	1.16	27.46	0	1.16	2.94	38.25	0	69.74
Published Freq	6.70%	18(Asi 16)	0.1	1.8	18(Afr 6.3)	0.1(Afr0.01)	0.6	1.7	34(As1 18)	Asi 0.7	53(Asi 65)

FIGURE 87

from SNP Consortium

TSC_NO	TWT_NO	Allele Frequency AfroAmerican	Allele Frequency Asian	Allele Frequency Caucasian
TSC0138091	96660	4.5	45	37.5
TSC0070363	63378	41	33	46
TSC0229600	65667	58	17	42
TSC0092788	55646	0	38	13
TSC0018629	37367	50	5	0
TSC0044098	55707	35	0	20
TSC0103676	64474	5	0	25
TSC0453079	64840	30	35	50
TSC0506473	41914	40	0	30
TSC0222745	25008	4	21	4.5
TSC0385435	17757	45	36	11
TSC1091265	124639	45	35.7	30
TSC0028346	43690	8.3	33	17
TSC0263586	100855	12.5	44	33
TSC0103744	69125	35	0	0
TSC0000358	56081	15	45	45
TSC0020498	56094	0	0	0
TSC0089797	35506	44	22	41

Each population comprised of 42 individuals

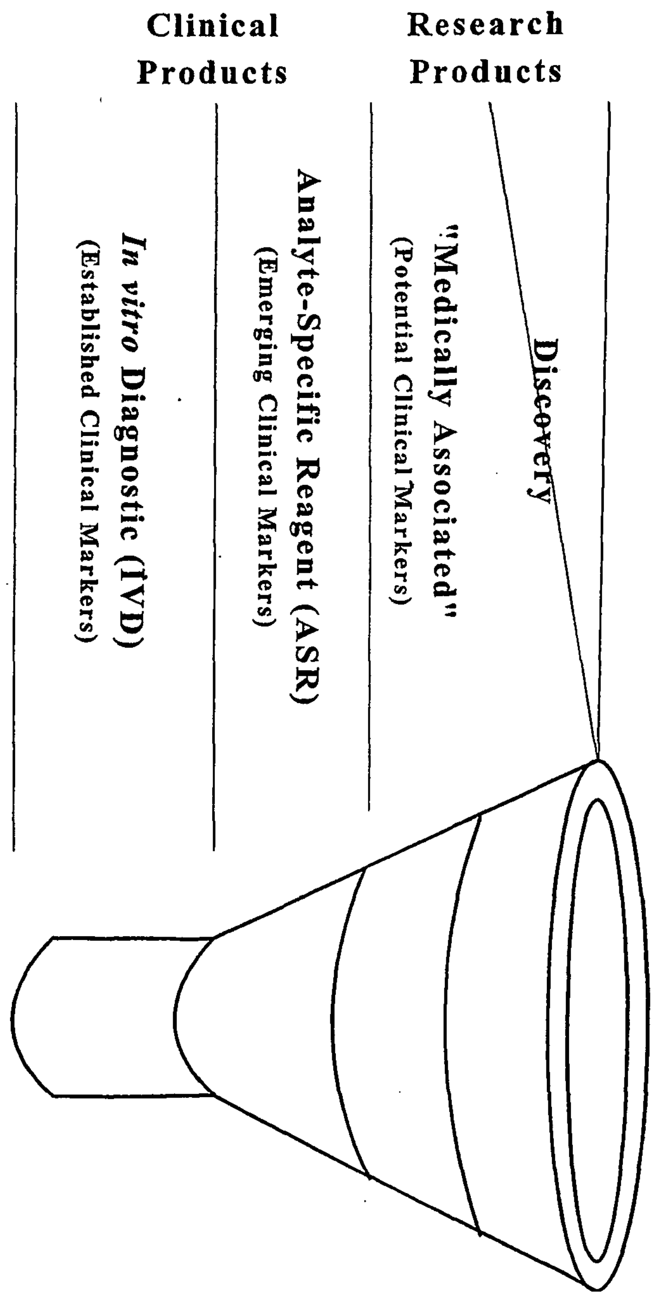

Research Products
Clinical Products
Discovery
"Medically Associated"
(Potential Clinical Markers)
Analyte-Specific Reagent (ASR)
(Emerging Clinical Markers)
In vitro Diagnostic (IVD)
(Established Clinical Markers)

FIGURE 88

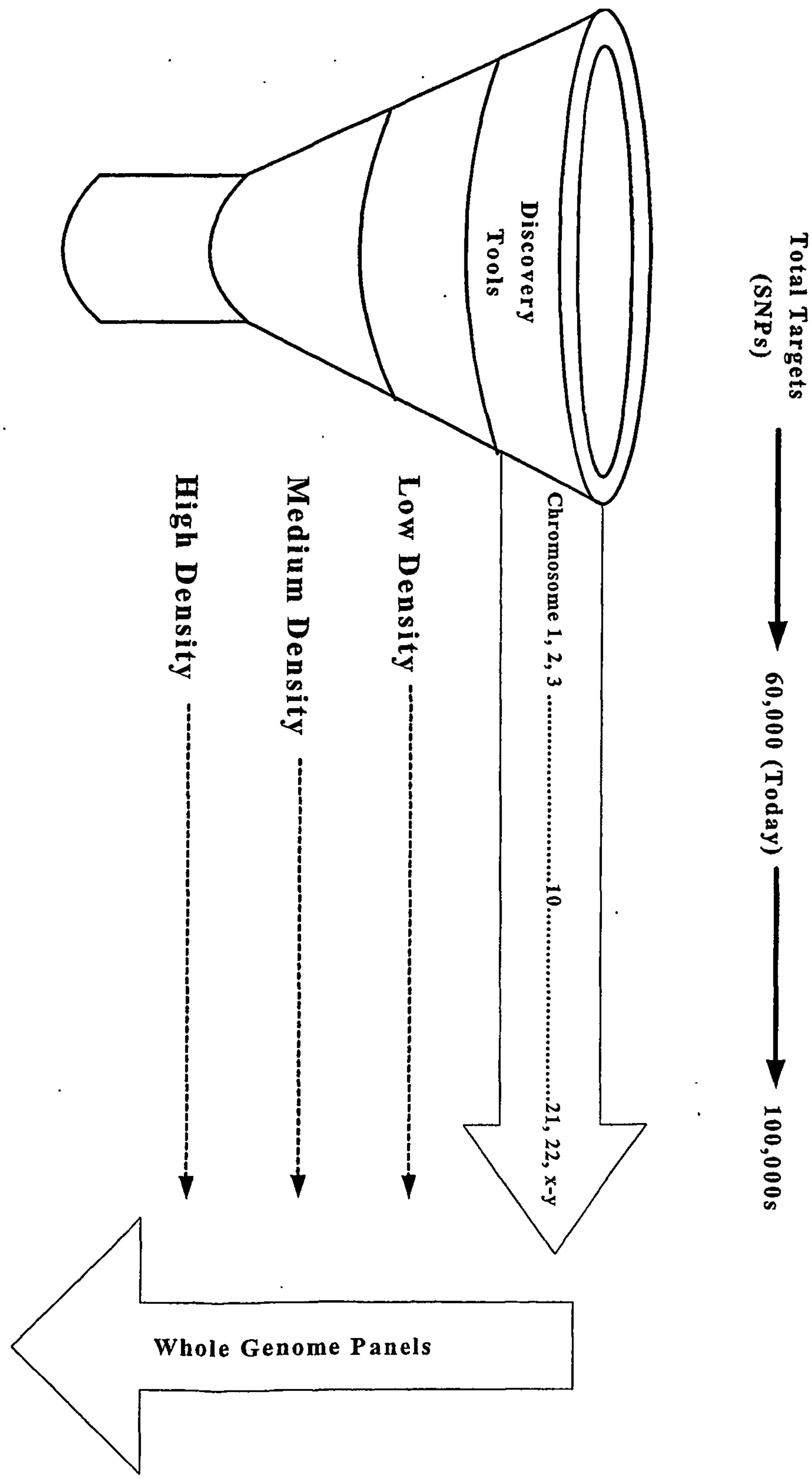

Total Targets (SNPs)
60,000 (Today)
100,000s
Discovery Tools
Chromosome 1, 2, 3 ....10..........21, 22, x-y
Low Density
Medium Density
High Density
Whole Genome Panels

FIGURE 89

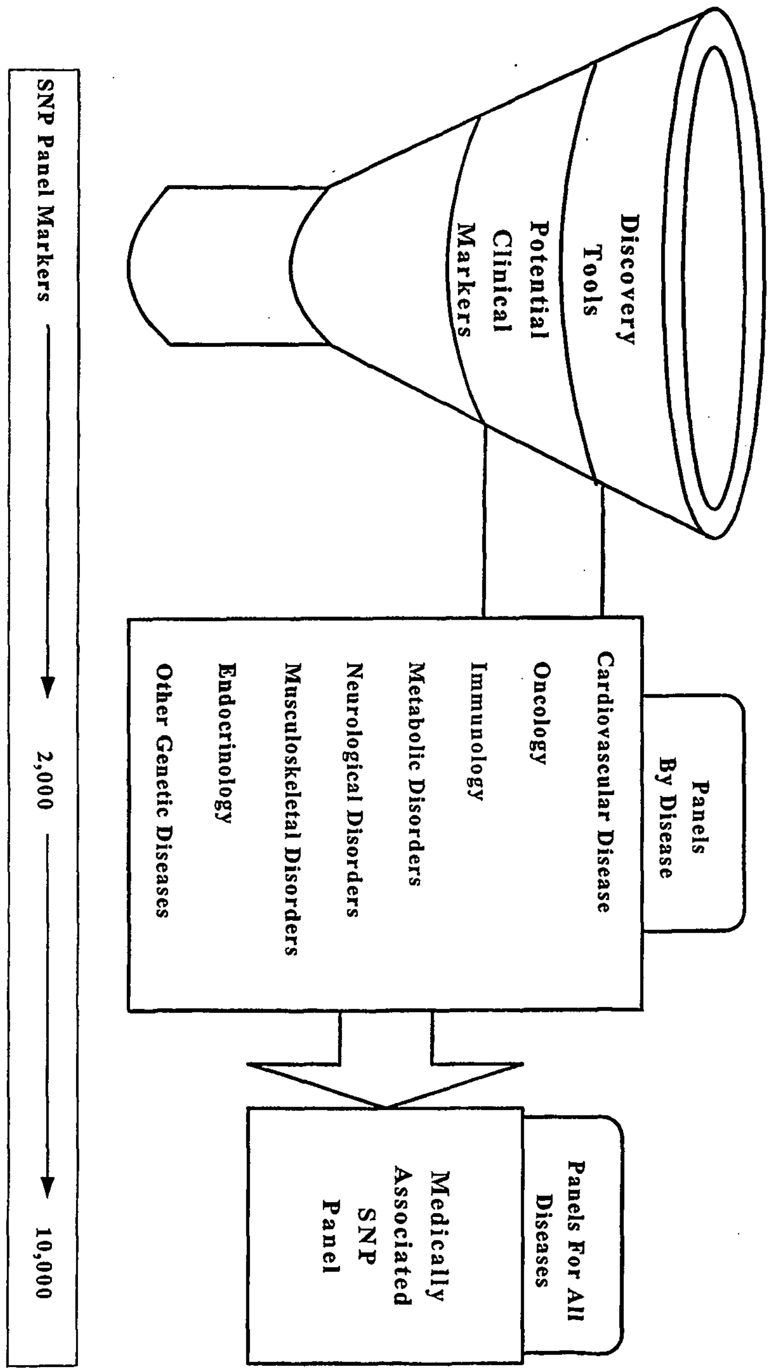
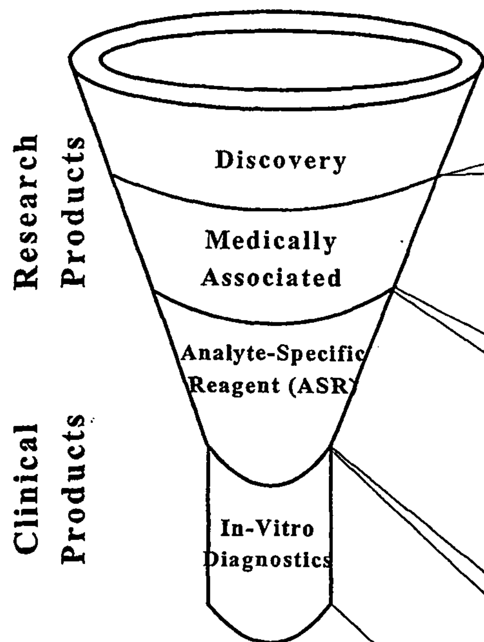

Discovery
Tools
Potential
Clinical
Markers
Panels
By Disease
Cardiovascular Disease
Oncology
Immunology
Metabolic Disorders
Neurological Disorders
Musculoskeletal Disorders
Endocrinology
Other Genetic Diseases
Panels For All
Diseases
Medically
Associated
SNP
Panel
SNP Panel Markers
2,000
10,000

FIGURE 90

FIGURE 91



	Genotyping/SNPs		Expression/RNA	
	Standard	Custom	Standard	Custom
Discovery	> 50K Unique SNP Assays Gene, Chromosome and Genome Panels	Requested SNP Assays Configured as Requested	cDNA Analysis All Human Genes	Requested RNA Analytes Configured for Research and/or HTS Applications as Requested
Medically Associated	Drug Metabolizing Enzyme Panel Disease Area Specific Panels Medically Associated Panel	Requested SNP Assays Configured as Requested	33 Analytes Available 100s Available Soon	Requested RNA Analytes Configured as Requested
Analyte-Specific Reagent (ASR)	ASRs Launched ASRs under Development	Requested SNP Assays Configured for ASR Use	ASRs under Development	Requested RNA Analytes Configured for ASR Use
In-Vitro Diagnostics	Submissions	Requested SNP Assays Configured for IVD Use	Submissions	Requested RNA Analytes Configured for IVD Use

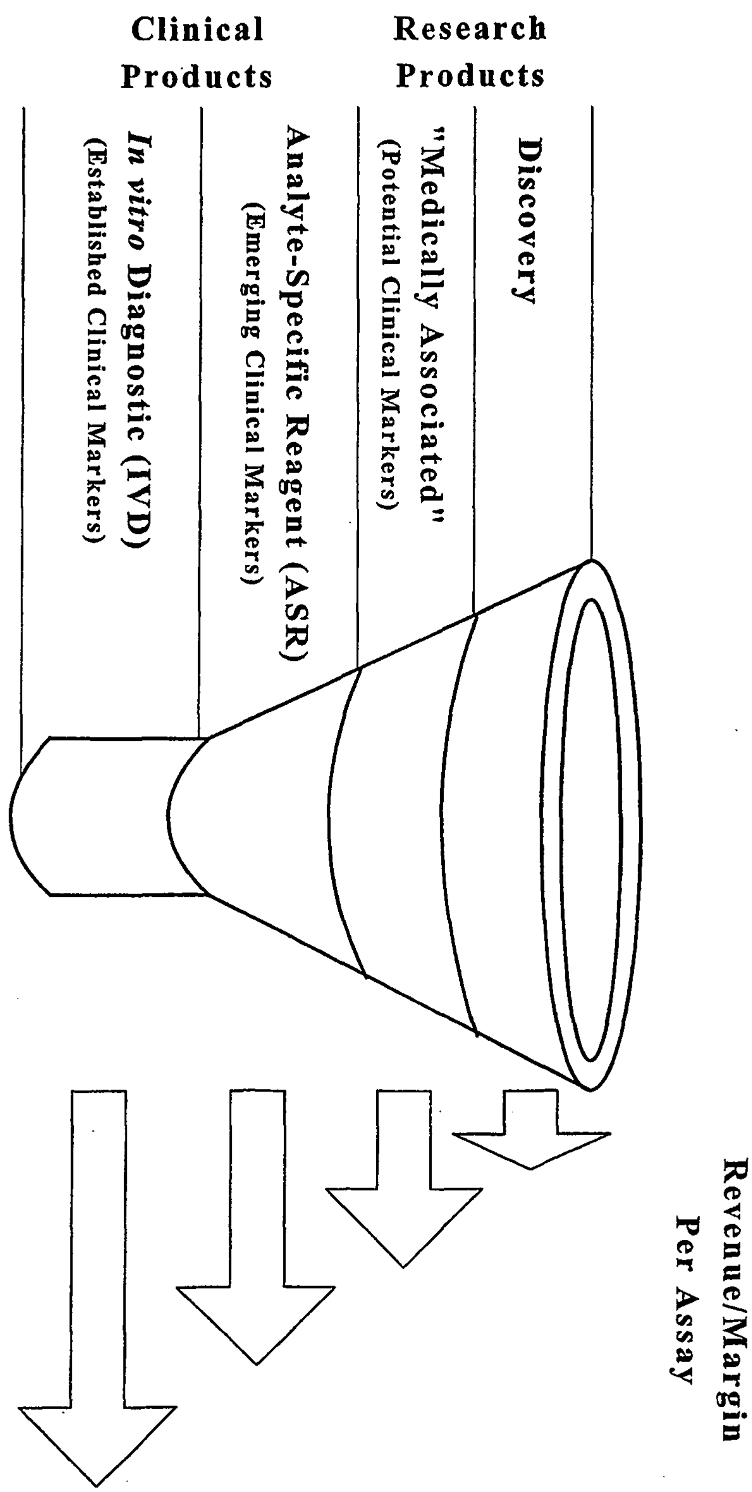

Research Products
Clinical Products
Discovery
"Medically Associated"
(Potential Clinical Markers)
Analyte-Specific Reagent (ASR)
(Emerging Clinical Markers)
In vitro Diagnostic (IVD)
(Established Clinical Markers)
Revenue/Margin Per Assay

FIGURE 92

FIGURE 93

Disease Associated Assay Development Process

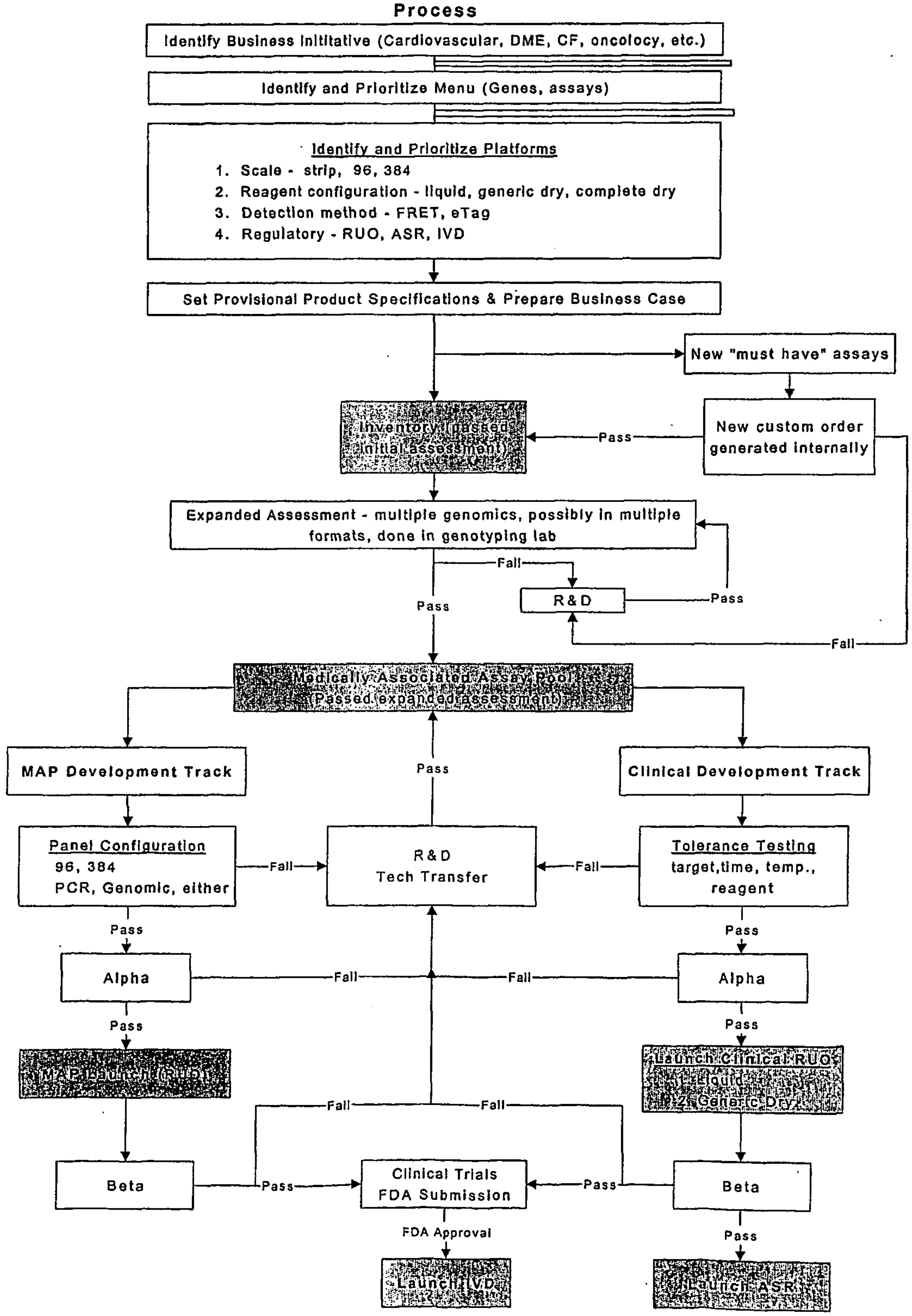


FIGURE 94

ASR Fast Track Overview

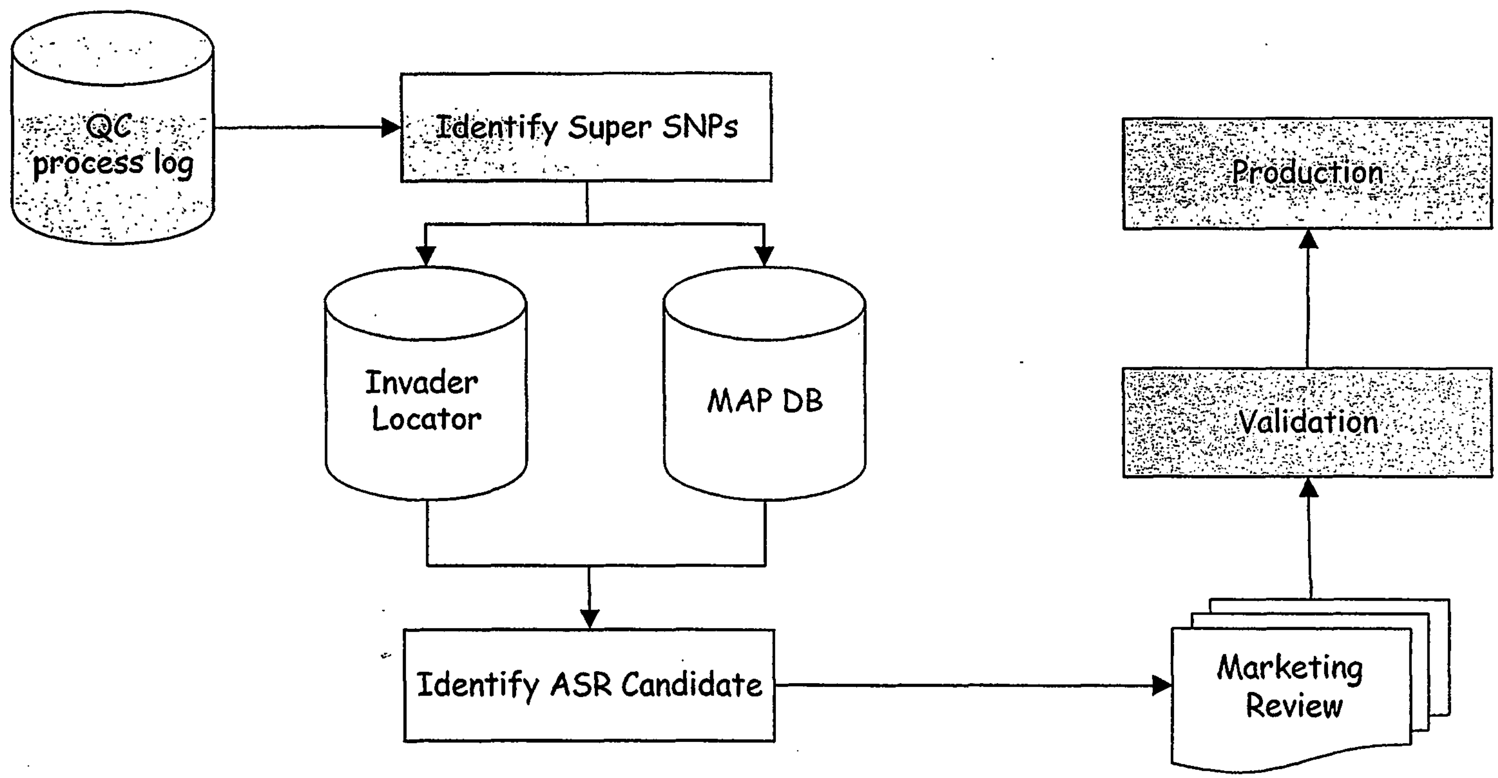


FIGURE 95

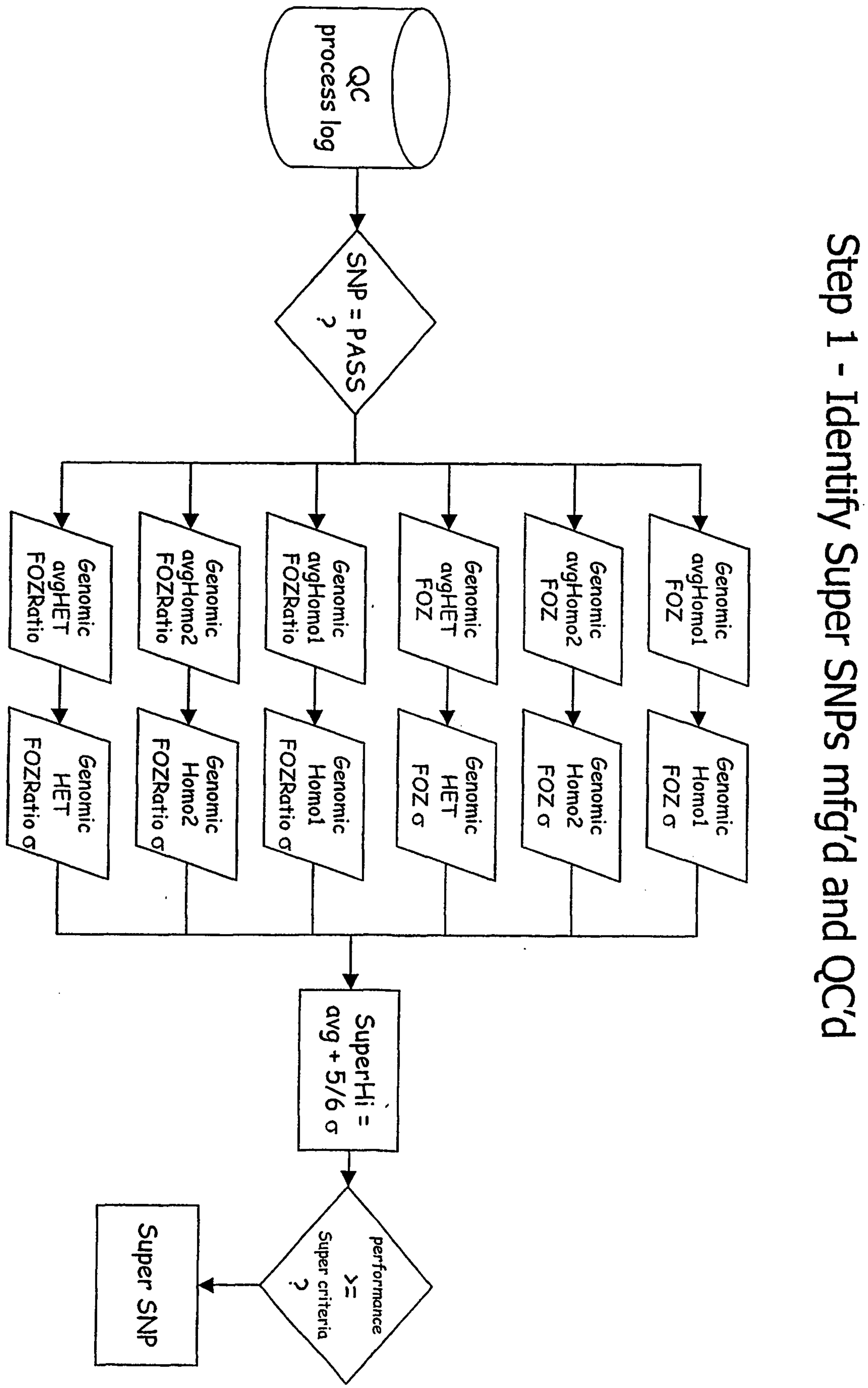

Step 1 - Identify Super SNPs mfg'd and QC'd
QC process log
SNP = PASS ?
Genomic avgHomo1 FOZ
Genomic avgHomo2 FOZ
Genomic avgHET FOZ
Genomic avgHomo1 FOZRatio
Genomic avgHomo2 FOZRatio
Genomic avgHET FOZRatio
Genomic Homo1 FOZ σ
Genomic Homo2 FOZ σ
Genomic HET FOZ σ
Genomic Homo1 FOZRatio σ
Genomic Homo2 FOZRatio σ
Genomic HET FOZRatio σ
SuperHi = avg + 5/6 σ
performance >= Super criteria ?
Super SNP

FIGURE 96

Invader Assay	Oligo	Polymorphism	Sequence	Short Name
Factor V Leiden (1691) Monoplex	Invader		5'-TAATCTGTAAGAGCAGATCCCTGGACAGR*CC-3'	FV G1691A
	WT Probe	G	5'-AACGAGGCGCACGAGGAATACAGGTATTTTGTC-3'	FV G1691A
	Mut Probe	A	5'-AACGAGGCGCACAAGGAATACAGGTATTTTGTC-3'	FV G1691A
	FRET		5'-F**-CCTC-Q***-GTCTCGGTTTTCCGAGACGAGGGTGCGCCTCGTTT-3'	FV G1691A
	WT Target	G	5'-AAGGACAAAATACCTGTATTCCTCGCCTGTCCAGGGATCTGCTCTTACAGATTAGA-3'	FV G1691A
	Mut Target	A	5'-AAGGACAAAATACCTGTATTCCTTGCCTGTCCAGGGATCTGCTCTTACAGATTAGA-3'	FV G1691A
Factor II (Prothrombin) Monoplex	Invader		5'-TATGGTTCCCAATAAAAGTGACTCTCAGCT-3'	FII G20210A
	WT Probe	G	5'-AACGAGGCGCACGAGCCTCAATGCTCCC-3'	FII G20210A
	Mut Probe	A	5'-AACGAGGCGCACAAGCCTCAATGCTCCC-3'	FII G20210A
	FRET		5'-F**-CCTC-Q***-GTCTCGGTTTTCCGAGACGAGGGTGCGCCTCGTTT-3'	FII G20210A
	WT Target	G	5'-TAGCACTGGGAGCATTGAGGCTCGCTGAGAGTCACTTTTATTGGGAACCATAGTTTTAGAAACACAAAAAT -3'	FII G20210A
	Mut Target	A	5'- TAGCACTGGGAGCATTGAGGCTTGCTGAGAGTCACTTTTATTGGGAACCATAGTTTTAGAAACACAAAAAT -3'	FII G20210A
Methylenetetrahydrofolate reductase (677) Monoplex	Invader		5'-CAAAGAAAAGCTGCGTGATGATGAAATCGC -3'	MTHFR C677T
	WT Probe	C	5'-AACGAGGCGCACGCTCCCGCAGACACX-3'	MTHFR C677T
	Mut Probe	T	5'-AACGAGGCGCACACTCCCGCAGACACCX -3'	MTHFR C677T
	FRET		5'-F**-CCTC-Q***-GTCTCGGTTTTCCGAGACGAGGGTGCGCCTCGTTT-3'	MTHFR C677T
	WT Target	C	5'- GAAGGTGTCTGCGGGAGCCGATTTCATCATCACGCAGCTTTTCTTTGAGG -3'	MTHFR C677T
	Mut Target	T	5'- GAAGGTGTCTGCGGGAGTCGATTTCATCATCACGCAGCTTTTCTTTGAGG -3'	MTHFR C677T
Methylenetetrahydrofolate reductase (1298) Monoplex	Invader		5'-cccgagaggtaaagaacaaagacttcaaagacactta-3'	MTHFR A1298C
	WT Probe	C	5'-aacgaacgcgcagTcttcactggtcagctatt-3'	MTHFR A1298C
	Mut Probe	A	5'-aacgaacgcgcagGcttcactggtcagcatt-3'	MTHFR A1298C
	FRET		5'-FL-acgc(Cy3)gtctcggttttccgagacgcgtctgcgcgttcgU*U*U*-3' (U* = 2'-O-methyl U)	MTHFR A1298C
	WT Target	C	5' – ggagctgaccagtgaagAaagtgtctttgaagtctttgttctttacctctcgggattt-3'	MTHFR A1298C
	Mut Target	A	5' – ggagctgaccagtgaagCaagtgtctttgaagtctttgttctttacctctcgggattt-3'	MTHFR A1298C
Hereditary Hemochromatosis, C282Y, Monoplex	Invader		5'-cccctggggaagagcagagatatacgtc-3'	HH C282Y
	WT Probe	G	5'-aacgaacgcgcagGccaggtggagcattt-3'	HH C282Y
	Mut Probe	A	5'-aacgaacgcgcagAccaggtggagcac-3'	HH C282Y
	FRET		5'-FL-ctcc(Cy3)gtctcggttttccgagacggagctgcgcgttcgU*U*U*-3' (U* = 2'-O-methyl U)	HH C282Y
	WT Target		5'-ggtgctccacctggCacgtatatctctgctcttccccag-3'	HH C282Y
	Mut Target		5'-ggtgctccacctggTacgtatatctctgctcttccccag-3'	HH C282Y
Hereditary Hemochromatosis, H63D, Monoplex	Invader		5'-gggctccacacggcgactctcatt-3'	HH H63D
	WT Probe	C	5'-aagcacgcagcacGatcatagaacacgaacagttt-3'	HH H63D
	Mut Probe	G	5'-aagcacgcagcacCatcatagaacacgaacagttt-3'	HH H63D
	FRET		5'-FL-acgc(Cy3)gtctcggttttccgagacgcgtgtgctgcgtgcU*U*U*-3' (U* = 2'-O-methyl U)	HH H63D
	WT Target	C	5'-agctgttcgtgttctatgatCatgagagtcgccgtgtggagccccg-3'	HH H63D
	Mut Target	G	5'-agctgttcgtgttctatgatGatgagagtcgccgtgtggagccccg-3'	HH H63D
Factor V Leiden (G1691A) Biplex	Invader		5'- tctaatctgtaagagcagatccctggacagGcc-3'	FV G1691A
	Invader (1689)		5'- tctaatctgtaagagcagatccctggacagAcc-3'	FV G1691A
	WT Probe		5'-cgcgaggccgGaggaatacaggtattttgtcc-hexandiol-3'	FV G1691A
	Mut Probe		5' -aggccacggacgAaggaatacaggtattttgtc-hexandiol-3'	FV G1691A
	FRET (WT)		5'-Red Dye-tct-Z28-agccggttttccggctgagacggcctcgcg-hexandiol-3'	FV G1691A
	FRET (Mut)		5'-FAM-tct-Z28-agccggttttccggctgagacgtccgtggcct-hexandiol-3'	FV G1691A
	WT Target		5'-tcaaggacaaaatacctgtattcctCgcctgtccagggatctgctcttacagattagaagtgattt-3'	FV G1691A
	Mut Target		5'-tcaaggacaaaatacctgtattcctGgcctgtccagggatctgctcttacagattagaagtgattt-3'	FV G1691A
Factor II (G20210A) Biplex	Invader		5'-TATGGTTCCCAATAAAAGTGACTCTCAGCT-3'	FII G20210A
	WT Probe		5'-ACGGACGCGGAGGAGCCTCAATGCTACCAG-3'	FII G20210A
	Mut Probe		5'-AGGCCACGGACGAAGCCTCAATGCTCC-3'	FII G20210A
	FRET (WT)		5'-R*-TCT-Q***-TCGGCCTTTTGGCCGAGAGACTCCGCGTCCGT-3'	FII G20210A
	FRET (Mut)		5'-F**-TCT-Q***- AGCCGGTTTTCCGGCTGAGACGTCCGTGGCCT-3'	FII G20210A
	WT Target		5'-TAGCACTGGGAGCATTGAGGCTCGCTGAGAGTCACTTTTATTGGGAACCATAGTTTTAGAAACACAAAAAT -3'	FII G20210A
	Mut Target		5'-TAGCACTGGGAGCATTGAGGCTTGCTGAGAGTCACTTTTATTGGGAACCATAGTTTTAGAAACACAAAAAT -3'	FII G20210A
Methylenetetrahydrofolate reductase (677) Biplex	Invader		5'-CAAAGAAAAGCTGCGTGATGATGAAATCGC -3'	MTHFR C677T
	Invader (SIP)		5'-CAAAGAAAAGCTGCGTGATGATGAAATTGC -3'	MTHFR C677T
	WT Probe		5'-ACGGACGCGGAGGCTCCCGCAGACAC-3'	MTHFR C677T
	Mut Probe		5'-AGGCCACGGACGACTCCCGCAGACAC -3'	MTHFR C677T
	FRET (WT)		5'-F**-TCT-Q***-AGCCGGTTTTCCGGCTGAGACTCCGCGTCCGT-3'	MTHFR C677T
	FRET (Mut)		5'-R**-TCT-Q***-TCGGCCTTTTGGCCGAGAGACGTCCGTGGCCT-3'	MTHFR C677T
	WT Target		5'- GAAGGTGTCTGCGGGAGCCGATTTCATCATCACGCAGCTTTTCTTTGAGG -3'	MTHFR C677T

FIGURE 97A

Step 2 - Allele Call

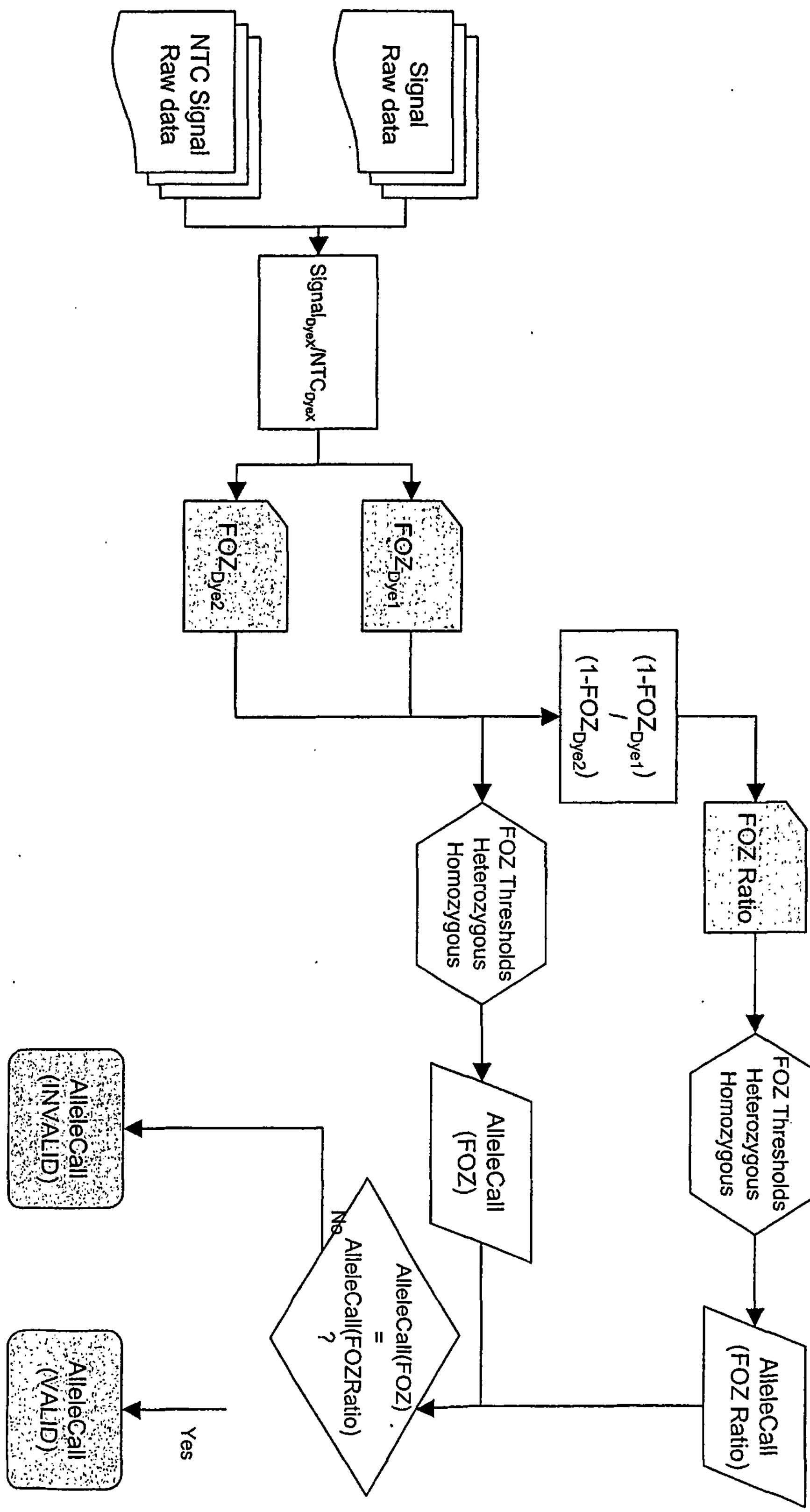


FIGURE 97B

Step 3 - Class Analysis

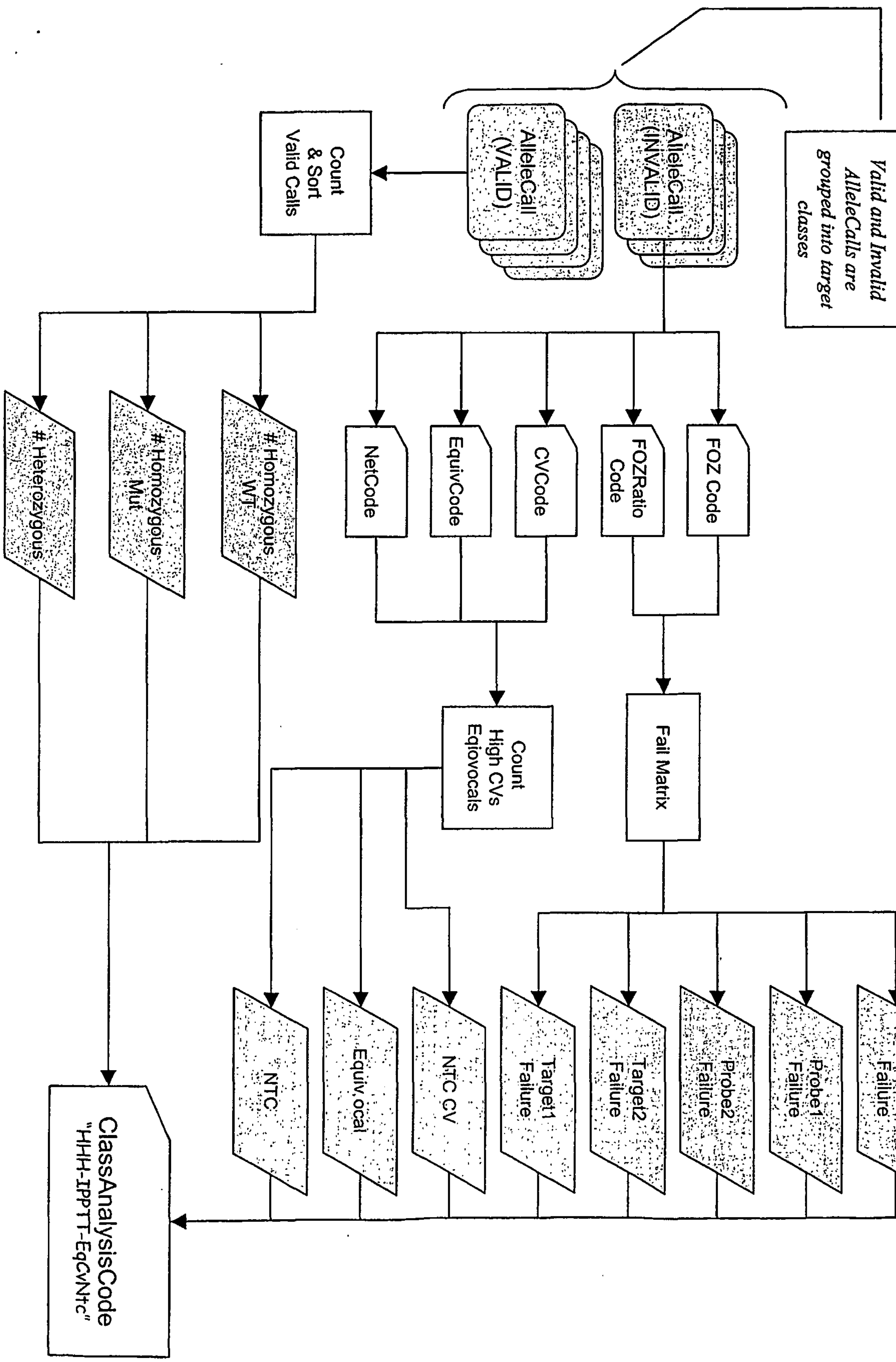


FIGURE 97C

Steps 4 - Target Class P/F/R Status

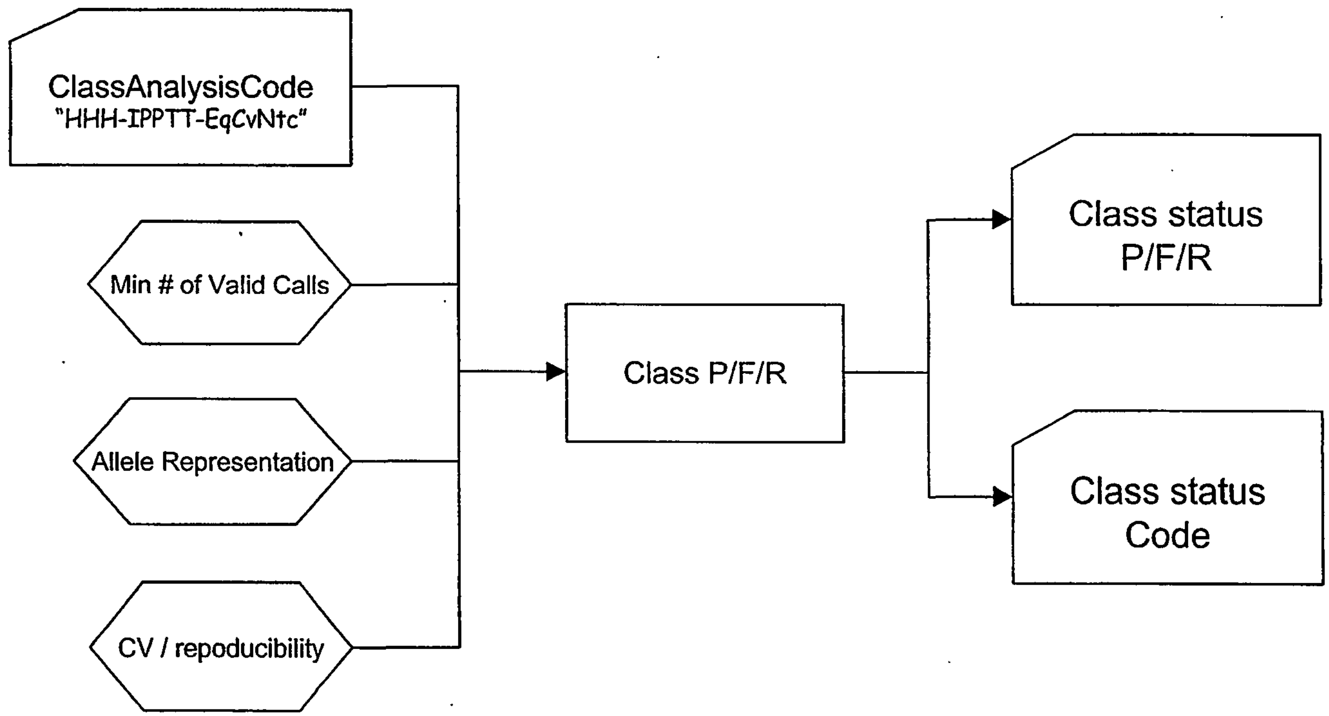
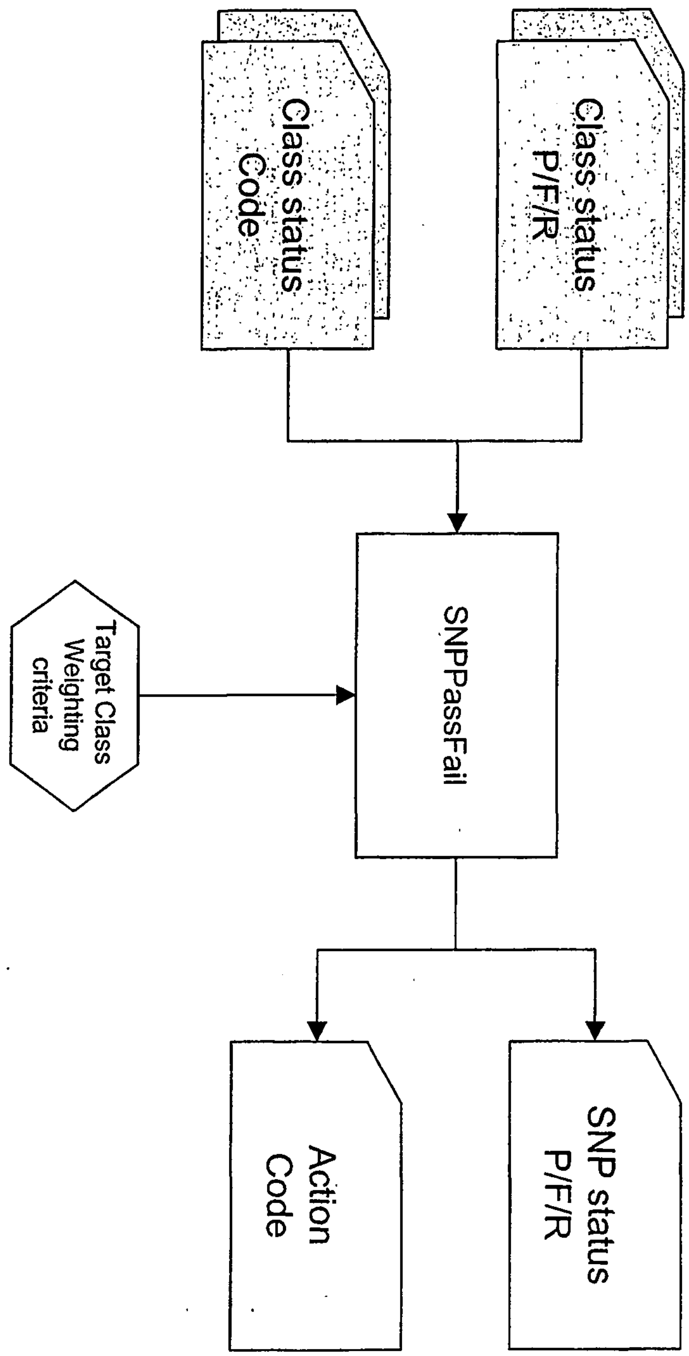


FIGURE 97D

Steps 5 & 6 - SNP Component ODN P/F/R Status and ODN Disposition



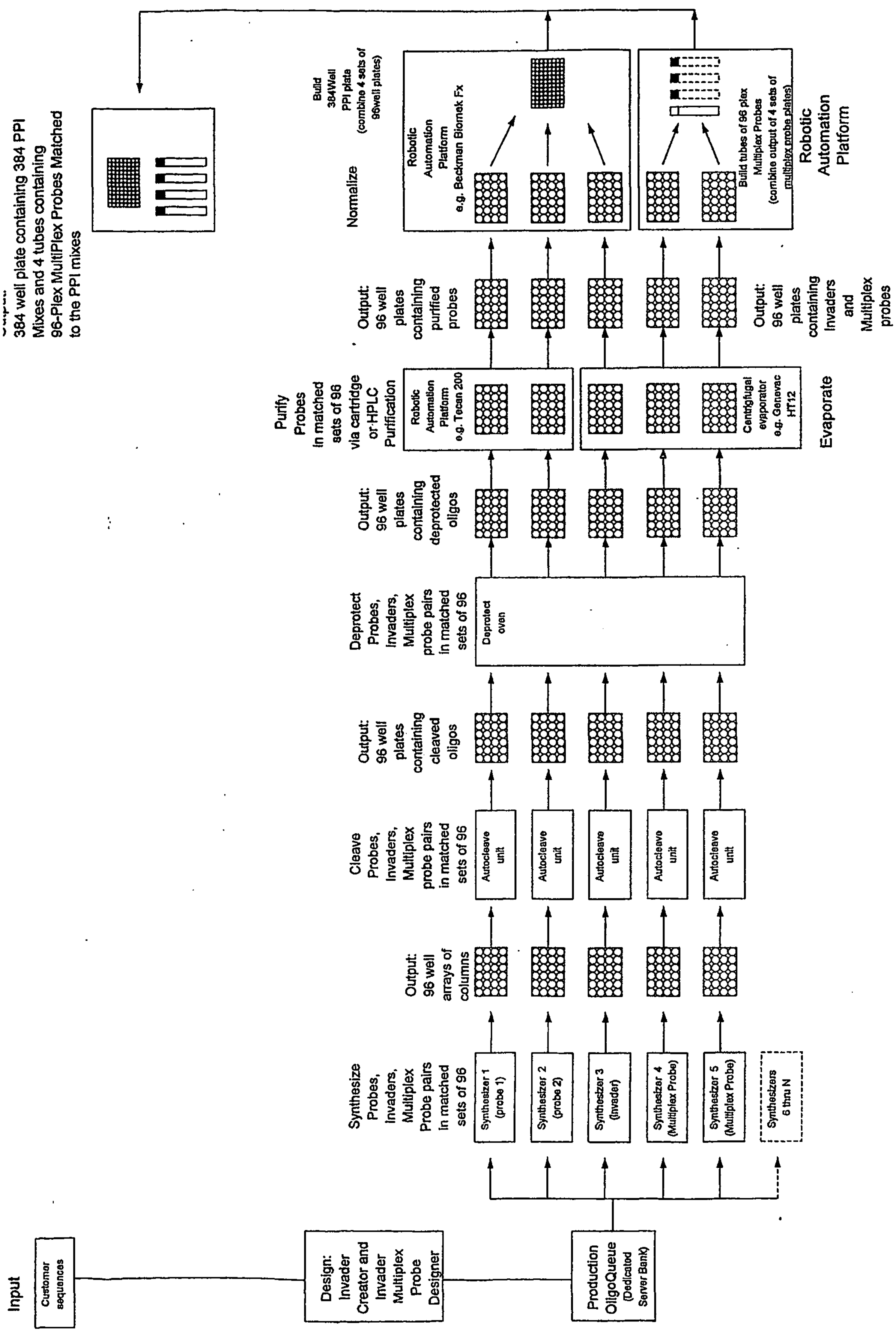

Input
Customer sequences
Design: Invader Creator and Invader Multiplex Probe Designer
Production OligoQueue (Dedicated Server Bank)
Synthesize Probes, Invaders, Multiplex Probe pairs in matched sets of 96
Synthesizer 1 (probe 1)
Synthesizer 2 (probe 2)
Synthesizer 3 (Invader)
Synthesizer 4 (Multiplex Probe)
Synthesizer 5 (Multiplex Probe)
Synthesizers 6 thru N
Output: 96 well arrays of columns
Cleave Probes, Invaders, Multiplex probe pairs in matched sets of 96
Autocleave unit
Output: 96 well plates containing cleaved oligos
Deprotect Probes, Invaders, Multiplex probe pairs in matched sets of 96
Deprotect oven
Output: 96 well plates containing deprotected oligos
Purify Probes in matched sets of 96 via cartridge or HPLC Purification
Robotic Automation Platform e.g. Tecan 200
Centrifugal evaporator e.g. Genevac HT12
Evaporate
Output: 96 well plates containing purified probes
Output: 96 well plates containing Invaders and Multiplex probes
Normalize
Robotic Automation Platform e.g. Beckman Biomek Fx
Build 384Well PPI plate (combine 4 sets of 96well plates)
Build tubes of 96 plex Multiplex Probes (combine output of 4 sets of multiplex probe plates)
Robotic Automation Platform
384 well plate containing 384 PPI Mixes and 4 tubes containing 96-Plex MultiPlex Probes Matched to the PPI mixes

FIGURE 98